Foreigner-Directed Speech and L2 Speech Learning in An Understudied Interactional Setting: The Case of Foreign-Domestic Helpers in Oman

Azza Al Kendi

Thesis submitted for the degree of Doctor of Philosophy (Integrated)
School of English Literature, Language and Linguistics

Newcastle University

February 2021

Abstract

Set in Arabic-speaking Oman, the present study investigates whether speech directed to foreign domestic helpers (FDH-directed speech) is modified when compared with speech addressed to native Arabic speakers. It also explores the FDH's ability to learn the sound system of their L2 in a near-naturalistic setting. In relation to input, the study explores whether there are any adaptations in native speakers' realizations of complex Arabic consonants, consonant clusters, and vowels in FDH-directed speech. By doing so, it compares the phonetic features of FDH-directed speech in relation to other speech registers such as foreigner-directed speech (FDS), infant-directed speech (IDS) and clear speech. The study also investigates whether foreign accentedness, religion and Arabic language experience, as indexed by length of residence (LoR), play a role in the extent of adaptations present in FDH-directed speech. In relation to L2 speech learning, the study investigates the extent to which FDHs are sensitive to the phonemic contrasts of Arabic and whether their production of complex Arabic consonants and consonant clusters is target-like. It also examines the social and linguistic factors (LoR, first and second language literacy) that play a role in the learnability of these sounds.

Speech recordings were collected from 22 Omani female native Arabic speakers who interacted 1) with their FDHs and 2) with a native-speaking adult (the order was reversed for half of the participants), in both instances using a spot the difference task. A picture naming task was then used to collect data for production data by the same FDHs, while perception data consisted of an AX forced choice task.

Results demonstrate the distinctiveness of FDH-directed speech from other speech registers. Neither simplification of complex sounds nor hyperarticulation of consonant contrasts were attested in FDH-directed speech, despite them being reported in other studies on FDS and IDS. We attribute this to the familiarity of the native speakers with their FDHs and the formulaic nature of their daily interactions. Expansion of vowel space was evident in this study, conforming with other FDS studies. Results from perception and production tasks revealed that FDHs fell short of native-like performance, despite the more naturalistic setting and regardless of LoR. L1 and L2 literacy played varying roles in FDHs' phonological sensitivity and production of certain contrasts. The study is original is terms of showing that FDS is not an

automatic outcome of interactions with L2 speakers and links these results with the unusual social setting.

Acknowledgements

Gratefulness and gratitude go first to Almighty God who had bestowed upon me strength, patience, and health during the course of my work. I was able to confidently accomplish my work despite all the odds, strongly believing in His words “And whatever of blessings and good things you have, it is from Allah; thereafter when adversity touches you, then Him you implore.”

After God, my sincerest and uttermost gratitude go to my supervisors Ghada Khattab and Martha Young-Scholten for the time and effort they spent to make sure my thesis is up to the standards. My deepest appreciation goes to my first supervisor Ghada Khattab for her invaluable ideas, guidance, and encouragement which had always raised my confidence and improved my work. She never was reluctant to meet with me whenever I asked for it and had always been a beacon of light guiding me through my struggles. Thanks also go to my second supervisor Martha Young-Scholten for encouraging me to investigate a group of vulnerable second language learners who have been ignored by the wider second language acquisition research and for her valuable comments on my thesis.

My deepest gratitude also goes to my husband, Younis, for believing in me and supporting me throughout my study at Newcastle. I cannot imagine myself surviving the Ph.D. life as a mom without him. His compromises were the most appreciated. I also thank my three beloved children for being there with me throughout my stay at Newcastle, enlightening my life and adding lots of fun to it. We survived those dark desperate days of winter together. I have learned that doing a Ph.D. can be a lonely process. However, I have never felt lonely as I was always surrounded by the voices of my little ones. I am so lucky to have you!

This Ph.D. thesis would not have been possible without the help and support of a number of people. First, I am indebted to Dr. Jalal Al-Tamimi for teaching me how to write Praat scripts and for his continuous feedback on scripts I designed. I am also grateful to him for providing me with feedback on my statistical analyses. I am also thankful to Dr. Danielle Turton and Dr. Lauren Ackerman for teaching me the basics of R software and providing me with feedback on some of my statistical analyses. My thanks also go out for my dearest friend Dr. Suaad Ambu Saidi for her friendship and for always being supportive of my work.

I would also like to say a heartfelt thank you to my family, my mom, my dad, my siblings and my mother in law, for their prayers and support throughout my Ph.D. journey. Special thanks go to my mom and my sister, Huda, for helping me reach out for participants in my study. Collecting data for this study would not have been possible without their help.

Table of Contents

List of Tables

List of Figures

Abbreviations

SLA: second language acquisition

FDH: foreign domestic helper

AoA: age of arrival

LoR: length of residence

AO: age of onset

L1: first language

L2: second language

NS: native speaker

NNS: non-native speaker

L2er: second language learner

FDS: foreigner directed speech

ADS: adult directed speech

IDS: infant directed speech

MSA: modern standard Arabic

CAT: communication accommodation theory

CPH: critical period hypothesis

H & H: hyper and hypo theory

VOT: voice onset time

CoG: center of gravity

LMEM: linear mixed effect model

GLMM: generalized linear mixed model

M1: Center of gravity

M2: Standard deviation

M3: Skewness

M4: Kurtosis

Chapter 1. Introduction

1.0 Area and Scope

Over the past decades, a great deal of second language acquisition (SLA) and applied linguistics research has focused on the factors that play a role in successful second language acquisition (L2A). Among the factors investigated (e.g. age, acculturation, aptitude), consideration of input expands the scope of SLA from a focus on interlanguage production as a demonstration of the L2 learners' underlying processes to include "a concern for the learner's linguistic environment and its role in facilitating these processes" (Pica *et al.* 1986). This interest was primarily motivated by work on first language (L1) acquisition that looks at how oral input which the speaker adjusted to the learner's language development level might aid language acquisition to take place (Richards and Gallaway 1993), and earlier work in SLA on 'foreigner talk' (Ferguson 1971, 1975). 'Foreigner' in this context refers to people who are assumed to have less knowledge of the language than the native speaker. Of particular concern in this area were the modifications made to spoken language, and subsequently, to conversational speech patterns in native speaker (NS)-non-native speaker (NNS) interactions in and outside the classroom (Wesche 1994). A great deal of research now exists that documents the nature of these modifications, their functions, the conditions under which they take place and whether a link between the context/interaction features and acquisition outcomes exists (e.g. Ferguson 1971, 1975; Freed 1981; Chaudron 1988; Hatch 1983; Long 1981; Long and Porter 1985; Pica *et al.* 1986; Uther *et al.* 2007; Knoll *et al.* 2015; Alfallaj 2016; Zuraida and Fitri 2019, see Chapter 3).

Many studies investigating foreigner talk have focused on its morpho-syntactic and discourse/conversational aspects (Ferguson 1971, 1975; Freed 1981; Pica *et al.* 1986; Gass and Mackey 2007; Gass 2009; Alfallaj 2016; Zuraida and Fitri 2019). Phonological aspects were also investigated but were limited in their focus or were not based on instrumental analyses (Katz 1977; Katz 1981; Clyne 1981; Hatch 1983). Only a few subsequent studies have investigated the acoustic and prosodic characteristics of speech directed to foreigners (foreigner-directed speech hence after) (Biersack *et al.* 2005; Knoll and Scharrer 2007; Scarborough *et al.* 2007; Uther *et al.* 2007; Hazan *et al.* 2015; Knoll *et al.* 2015). Foreigner-directed speech (FDS) studies have concentrated on investigating features that have been investigated in infant-directed speech (IDS) and clear speech including vowel space area, vowel length, fundamental frequency (F0) and pitch contours (Chapter 3). This was due to the assumption that foreigners are comparable to infants or hearing-impaired listeners with regards to their limited access to the language code and hence might benefit from modified L2 input (Bradlow and Bent 2002; Uther *et al.* 2007). Despite some interesting results reported about FDS (e.g. expanded vowel space and slower rate, see Chapter 3), this area of research has remained understudied. The present study was motivated by the needs to further investigate potential modifications in speech addressed to foreigners who have not, to my best knowledge, been studied at all in this kind of research and to expand research on this area by exploring phonetic variables that have not been examined in this line of research. Additionally, this study was motivated by the need to explore these adjustments in a setting that is more naturalistic and less ephemeral than brief encounters between NSs and foreigners, which make up most of the researched contexts. The setting chosen is that of an NS employer-foreign domestic helper (FDH) interaction. In order to understand why this setting was especially interesting for the current study, Section 1.1 will provide a discussion of the social and linguistic situation of adult migrants in the Middle East.

Studies that investigated foreigner talk in the 1970s and 1980s or FDS more recently focus primarily on the input addressed to foreigners. The current study also examines FDHs' L2 speech production and perception. SLA studies highlight the difficulty adult L2 learners face when learning L2 speech (Flege 1981; Best and Strange 1992; Akahane-Yamada 1995; Flege *et al.* 1995; Iverson and Khul 1996). While a naturalistic setting can be advantageous for children of immigrants, who also receive formal instruction in the L2 country, adult immigrants may be disadvantaged due to difficulty in getting access to formal instruction and/or coming from low educational backgrounds (Klein and Perdue 1997; Craats *et al.* 2006). As will be further discussed in Section 1.2, research on oral/aural L2 performance of low educated learners is scarce. Studies investigating FDHs' L2 performance has especially focused on features which have prompted researchers to claim that a pidgin variety develops in these contexts, but this research has mostly focused on morphosyntactic features, as will be discussed in Section 1.1. The present study was motivated to explore FDHs' speech learning as well as to examine the potential factors playing a role in their L2 learning. The distinction between acquisition and learning is not of paramount importance in the present study and the two terms might be used interchangeably in some of the discussions in this thesis. Nonetheless, it should be pointed out that some researchers, especially in applied linguistics, draw a clear distinction between the two terms with respect to syntax and morphosyntax. Krashen (1982) assumes that *acquisition* is an unconscious process while *learning* requires the learner to be aware of the rules and develop metalinguistic knowledge of the language. Fundamentally, the former concept is a concomitant of L1 acquisition, whereas the latter is associated with L2 learning (Gee 1996). The term 'speech learning' was coined by Flege and Port (1981) to refer to the process in which L2 learners articulate and perceive speech sounds differently after extensive exposure to a foreign language. In support of this definition, James (1988: 30) affirms that learning new sounds entails learning "new patterns of articulation and perception". Hence, 'speech learning' provides two characteristics that are essential to the present study: perception and production.

1.1 Social and Linguistic Background of the Present Study's Learner Population

The Middle East has witnessed considerable and continuous labor migration as well as urbanization since the 1970s following the oil boom (Avram 2014; Bizri 2014). Foreign workers migrate into the area mostly from countries in South and South-East Asia and other parts of the Middle East, including Iran, India, Pakistan, Nepal, Bangladesh, Sri Lanka, Indonesia, Thailand and the Philippines (Bakir 2010; Avram 2014). Growing demand for live-in domestic helpers in the Middle East has involved the migration of thousands of female workers from these countries as well as from sub-Saharan Africa. These female workers are often initially employed on a two or three-year contract but can extend their stay with the same family or move to another house with a new contract (Bizri 2014). They can alternatively look for a new job in another country in the Middle East. As part of their contract, FDHs are provided with monthly salaries along with accommodation, food, and clothes. As more families have started employing FDHs with the increase in means, 'hiring' FDHs, as it is often referred to, has become a way of securing more prestige and high status (Al-Najjar 2004). FDHs are expected to work for long hours during the day for minimal pay (Al-Najjar 2004; Godfrey *et al.* 2004).

Oman, the setting of the present study, is one of the countries in the Middle East that has registered high numbers of FDHs in households. Under the leadership of His Majesty Sultan Qaboos bin Said, Oman's GDP rose and this has led to an increasing number of migrant workers including professionals (e.g. educators, technical experts, highly skilled workers) and low skilled workers among whom are FDHs as well as gardeners, cleaners and so on (Ghosal and Porkodi 2014). The first FDHs who migrated to Oman were mostly South Asians. In addition, FDHs from the Philippines, Sri Lanka, and Indonesia were also seen in many households. More recently, FDHs come from other regions from various African countries.

In the presence of this large-scale influx of migrant workers, a drastic change in the linguistic situation has taken place in the Middle East generally, and the Arabian Gulf states specifically.[1] In Gulf cities, urbanization has established prevalent use of contact languages such as English, Hindi, Urdu as well as several Arabic dialects spoken by the local Arabs or Arab migrants (Miller 2004). Furthermore, with the diverse nationalities of the foreign workers, the first languages of these workers are typologically and genetically diverse. According to Avram (2014), in the United Arab Emirates, there are around 150 ethnic groups who speak around 100 languages. FDHs in the Gulf, including in Oman, are usually addressed either in Arabic or in English (or both) [2].

FDHs acquire Arabic in a naturalistic setting through their interactions mainly with members of their host family but also with other non-native speakers. They do not get the chance to learn Arabic in a classroom setting where they would explicitly learn the pronunciation and forms of the L2, and hardly get any feedback on their production errors.[3] Instead, they pick up the language from the ambient input, in some way similarly to how children acquire their native language, though marked differences in the speech addressed to each group are expected. Arabic involves diverse varieties as well as Modern Standard Arabic (MSA) which is only spoken on formal occasions. Most FDHs have a basic education in their first languages but some have no formal schooling (Bizri 2014).

One consequence for the presence of a heterogeneous group of migrant workers who lack knowledge of the target language is the use of special foreigner talk by the native speakers of that community when addressing foreigners. For example, Bizri (2009; 2014) identifies what she calls "Pidgin Madam" between Lebanese female employers and their Sri-Lankan FDHs. She observes that the form of Arabic that was used by (and to) FDHs in Lebanon was different from the conventional Lebanese Arabic used by local Lebanese Arabs. The most fluent speakers of this form of Arabic, other than the FDHs themselves were the female employers

---

[1] Arabic Gulf States include Bahrain, Saudi Arabia, United Arab Emirates, Oman, Qatar and Kuwait.

[2] Some FDHs who are fluent speakers of English (e.g. Filipinos) are addressed in English especially when members of their household are fluent speakers of English. According to Avram (2014), English is not widely spoken, neither by immigrant foreigners from South Asia except for those from the Philippines nor by the local Arab population in the Gulf countries.

[3] Based on my own observation, Filipino domestic helpers get to learn some Arabic words as part of a training programme on domestic help that they receive before migration.

and their children (ibid). The description Bizri (2009) provides about this variety was the most significant morphological and syntactic features. Dashti (2013), likewise, demonstrates the use of simplified syntax by native speakers of Kuwaiti Arabic when addressing their FDHs from various Asian countries.

1.2 Significance of the Present Study

The literature on the linguistic experience of foreigners in the Arabic-speaking World has been mostly discussed from a sociolinguistic point of view, either discussing speech addressed to foreigners (e.g. Smart 1990; Bizri 2009; Dashti 2013) or the kind of Pidgin Arabic spoken among the foreigners themselves (e.g. Smart 1990; Bakir 2010; Alghamdi 2014; Almoaily 2014; Al-Zubeiry 2015). Examining the learning experience of these foreigners from an SLA point of view is understudied. Evidence for use of FDS by Arab native speakers when addressing foreigners is scarce and only for morpho-syntax (e.g. Dashti 2013; Bizri 2009). Phonological and acoustic descriptions of FDS specifically used with migrant workers do not exist. The only description of the phonology of foreigner talk is Smart's (1990) brief report, which was not based on any empirical research. Smart observes that Arabs preserved Gulf Arabic phonology when conversing with foreigners (See Chapter 5).

Crucially, the present study contributes to the understudied research area on FDS. Studies so far conducted on FDS have been predominantly concerned with investigating acoustic properties of input to strange (as opposed to familiar) foreign accented learners and compared these to IDS and clear speech with a focus on vowel hyper-articulation, speech rate and pitch (e.g. Biersack *et al.* 2005; Uther *et al.* 2007). In addition to examining features currently reported in FDS, this study takes a further step by exploring two other aspects: 1) NSs' realization of complex Arabic consonants when addressing foreigners and whether consonantal contrasts are enhanced in this speech style, and 2) whether any adaptations in FDS vary based on the foreigners' length of residence in the second language community (LoR) as is typically found in SLA studies or speech to children of different ages.

The current study contributes to the wider SLA research community by examining the role of native language formal education and L2 literacy on adult L2 speech learning in a naturalistic acquisition setting. The vast majority of studies on SLA make use of convenience sampling,[4]

[4] Recruiting participants who are willing to volunteer and available.

and thus show an overreliance on a population which is WEIRD (Western, Educated, Industrialized, Rich, Democratic; Henrich *et al.* 2010). Much more research that has been published in journals including *Second Language Research, TESOL Quarterly,* and *Studies in Second Language Acquisition* rely on such samples (Bielow and Tarone 2004; Craats *et al.* 2006). Based on research conducted by Plonsky (2016), 67% of all samples used in 600 studies from six SLA journals involved college or university students. Furthermore, few studies include L1 education as a contributing variable (Craats *et al.* 2006; Young-Scholten 2013; 2018). This bias towards recruiting and examining highly educated L2 learners might have skewed our understanding of second language acquisition and led to under-explanation of how adults acquire a new language when we isolate factors such as native language literacy. Crucial differences between highly educated L2 learners and non- or low-literate learners have made this gap in research more apparent (Craat *et al.* 2006). Low socio-economic immigrants, generally, and low-educated immigrants, particularly, seem more likely to face difficulty achieving reasonable levels of attainment in L2 oral proficiency and may be more inclined to fossilize at an early stage of language development (Klein and Perdue 1997; Craats *et al.* 2006). To this end, the present study has the advantage of looking at whether FDHs' educational level influences their L2 performance, doing so with respect to Arabic, a language that has been largely understudied in SLA research.

1.3 Focus and Aims

The current study examines the speech learning of Arabic and to a greater extent the linguistic properties of NS input in an NS-FDH interaction setting where learners had no formal instruction. Specifically, it examines adaptations in the production of complex Arabic consonants (/tˤ, ðˤ, sˤ, χ, ʁ, q, ħ, ʕ, θ, ð /), vowels (/aː/, /iː/ and /uː/) and two-member consonant clusters in FDH-directed speech by comparing them to those present in speech directed to an NS control (henceforth adult-directed speech (ADS)). Whether adaptations occur will be discussed in relation to previous work on IDS, FDS and clear speech as well as within two main theoretical frameworks: Communication Accommodation Theory and Hyper and Hypo theory (See Chapter 3). It also examines whether Arabic language experience indexed by LoR plays a role in any adaptations attested in FDH-directed speech.

On a smaller scale, the present study also examines FDHs' speech learning by focusing on their phonological sensitivity to Arabic phonemic contrasts as well as their production accuracy of complex Arabic consonants and consonant clusters. Most importantly, it explores the social and cognitive factors (LoR, L1 schooling and L2 literacy) that play a role in the foreigners' learnability of the previous aspects by controlling for their first language (L1) and age of arrival (AoA).

The following section will present the research questions the study seeks to answer.

1.4 Research Questions

The present study aims to answer the following questions:

Q1) What (if any) systematic adaptations are made in consonant, consonant cluster and vowel productions in FDH-directed speech? To what extent are these adaptations similar to or different from those reported in IDS, FDS, and clear speech research?

Q1a) Does FDH-directed speech contain hyperarticulation/enhancement of sound contrasts?

Q2b) Do variations in FDH's foreign accentedness, LoR or religion play a role in any adaptations made in speech addressed to them?

Q2) To what extent are FDHs perceptually sensitive to the phonemic contrasts of their L2 Arabic? What social and cognitive factors (independent variables) are responsible for any variation present in the listeners' performance?

Q3) How accurate (or target-like) are FDHs at producing Arabic consonants and consonant clusters? What social and cognitive factors (independent variables) are responsible for any variation in the accuracy level of FDH's production skills?

1.5 Organization of the Thesis

In order to achieve the aims of this research and answer the research questions, this thesis is organized in the following way:

Chapter 2. This chapter will focus on a literature review of L2 speech learning. Theoretical models and studies concerning the perception of speech as well as the production of consonants and consonant clusters will be presented. It will also discuss studies on L2 learners' acquisition of Arabic. Additionally, it will present a literature review on the factors that affect L2A especially those that are relevant for the current study.

Chapter 3. This chapter will present the linguistic features of IDS, clear speech and FDS, and will discuss in more detail the phonological and acoustic features of these codes. Furthermore, the chapter will present the theoretical models that can be used to interpret the nature of native-speaker-non-native speaker interaction.

Chapter 4. A discussion of the sounds of interest will be presented in this chapter. It will also focus on the specific linguistic variables that will be investigated. It will provide a description of the acoustic and articulatory properties of the Arabic sounds of interest as well as a phonological description of consonant clusters in Arabic. This chapter will also present a comparison of the sound system inventories of the target language, Nizwa Arabic, with that of the L1s of the FDHs. It will also present the syllable shapes available in each language.

Chapter 5. This chapter will describe the hypotheses, the sample, the tools used to collect data, the data transcription protocols, the acoustic analysis and the statistical design for the current study.

Chapter 6. This chapter will present a small experiment that was set to determine foreign accent rating of FDHs. This rating was necessary to investigate its effect on hyperarticulation.

Chapter 7. Results regarding the realization of complex Arabic consonants, consonant clusters and vowels in FDH-directed speech will be presented in this chapter. Here, any phonological or acoustic adaptations present in FDH-directed speech will be revealed.

Chapter 8. This chapter will present results regarding the speech learning of FDHs. It will present results on their phonological sensitivity to phonemic contrasts as well as their production accuracy of consonants and consonant clusters. It will also cover results on the role of cognitive and social factors on the learners' performance.

Chapter 9. A general discussion of the findings obtained from the results will be provided in this chapter.

Chapter 9. This chapter will discuss the study's limitations and future directions before it provides its concluding remarks.

Chapter 2. Second Language Speech Learning

2.0 Introduction

This chapter will discuss literature on second language (L2) speech learning with a focus on speech perception and production. It will start with reviewing literature on L2 speech production. This will involve a section discussing foreign accent and segmental production and another section discussing production of L2 consonant clusters. This will be followed by a literature review on L2 speech perception. Then studies on L2 Arabic learning will be provided. After presenting the most prominent theoretical models and studies of speech perception and production, the chapter will provide a discussion of the factors that affect L2 speech learning and that are most relevant to the foreign domestic helper (FDH) context.

2.1 L2 Speech Production

2.1.1 Foreign accent and segmental production

A growing body of research has found that adults find it quite difficult to entirely get rid of a foreign accent, whereas most normal young children find it effortless to learn the sounds of their L1 and produce them without any trace of a foreign accent (Flege 1981; Moyer 2013; Huang 2014). Of all features of human language, accent is, perhaps, the most noticeable. A listener does not often need prior linguistic knowledge to detect a ‘foreign accent’, that is, pronunciation properties that are perceived as being unlike those of speakers of the native language (Thompson 1991). L2 speech production studies have focused on examining L2 speech production by looking at individual vowels and consonants, consonant clusters, words or full sentences (Leather and James 1996). The primary basis for a foreign accent is mispronunciations that result from segmental (but also suprasegmental) sound substitutions such as in Arabic-accented English production of *parking lot* with short-aspirated [p], leading it to be perceived as ‘barking lot’ (Flege 1981:445). In adult L2 learning, non-native patterns of pronunciation are extensive, potentially impacting a large number of segments in an inventory (Munro 1993). Additionally, a foreign accent is predictable in late L2 learners, even among those exposed to native speaker (NS) input for many years (Flege *et al.* 1995). Flege (1988) identifies a number of consequences that could result from speaking with a foreign accent, including reduced intelligibility, reduced acceptability, and negative evaluation. Likewise, Munro and Derwing (1995) propose that accentedness may lead listeners to exhibit prejudice such as discrimination, downgrading and irritation against some groups of speakers or accents. Moyer (2009), on the other hand, suggests that a potential advantage of an accent is that it could be perceived as an indicator of non-native proficiency, and thus lead to NSs' adjusting their speech input in interactions with L2 users. According to Varonis and Gass (1982), limited oral proficiency can trigger foreigner talk that is believed to enhance comprehension and communication.

Researchers have offered several explanations for foreign accents. A general agreement among these researchers is that the extent to which non-native L2 production is different from native production is greater for individuals who started learning the L2 in adolescence or adulthood relative to young starters (Long 1990). Young children under the age of 5-6 years old, unlike adults, could learn new languages with no trace of a foreign accent (Flege 1987). Given the comparative ease with which children learn to produce the sounds of their L1 with the appropriate phonetic specificities, they seem to have recognized the properties of the ambient input around them (Best 1992). This has motivated the assumption that L2 speech production is constrained by a 'critical period' resulting from the loss of neural plasticity (Patkowski 1989). In his proposal of the critical period hypothesis (CPH), Lennberg (1967) claims that the ability to learn a new language declines after pubetry due to the completion of neural hemispheric lateralization in the brain. Scovel (1988) claims that phonology more than any other linguistic ability is the most influenced by the age of first exposure to a second language. Scovel (1988: 62) also postulates that of all aspects of language, L2 speech production is the most affected by the critical period because "pronunciation requires an incredible talent for sensory feedback of where the articulators are and what they are doing. All other aspects of language are entirely 'cognitive' or 'perceptual' in that they have no physical reality".

Studies involving naturalistic exposure to L2 English appear to support the existence of a critical period (Oyama 1975; Seliger *et al.* 1975; Seliger 1975; Scovel 1981; Fathman 1982; Flege *et al.* 1995). However, informal observation and empirical research have also revealed exceptions to this hypothesis. L2 learners are sometimes able to achieve native-like proficiency in a second language (see Ioup *et al.* 1994). Despite the existence of such exceptions, the CPH should not be simply invalidated due to the amassed evidence available in support for it (Birdsong 1999).

Oyama (1976) investigates whether native-like mastery of L2 phonology can be achieved regardless of the age of acquisition in a study of 60 native Italian male immigrants in America.

The subjects varied on their age of arrival (AoA) to the L2 country (6-20 years) as well as their length of residence (LoR) (5-18 years). American-English native listeners judged the learners' performance on reading-aloud and storytelling tasks. The results revealed that the learners' performance was strongly correlated with their AoA but not LoR. Oyama concludes that his findings support the existence of a biologically determined decline that limit the ultimate acquisition of L2 phonology. Similarly, Flege *et al.* (1995) finds in a study of 240 native-Italian-adult-immigrants in Canada that those who arrived to the L2 community early in life were perceived more native-like in their production of English consonants compared to those who arrived later in life.

Despite the general belief on the existence of a critical period, it is still unknown what exactly causes a foreign accent (Flege 1995). Flege argues that it is still unknown whether the effect of age of learning (AoL) on the degree of a foreign accent is the result of "neurofunctional reorganization", cognitive alternations, or sociolinguistic factors. In addition, it is hardly known whether errors in production have a motoric basis or whether they are a result of an inability to perceptually discriminate the relevant phonetic properties of the L2 phonemes. Similarly, Moyer (2009) argues that AoA (for immigrants) or generally age of onset (AO) is fundamentally related to other factors such as instruction in the target language, length of residence (LoR) and socio-psychological factors (e.g. attitude, motivation, etc.). This suggests that much of the evidence in support of a critical period has been over-interpreted unless factors other than AoL or AO have been taken into consideration (Moyer 2009). The variation in these interpretations alludes to the complexity of understanding the causes of a foreign accent (Flege 1995). Section 2.4 will shed more light on some of these factors in more detail

Proponents of the critical period postulate that segmental production and perception might be affected differently by the critical period (Scovel 1988; Flege 1995). This brings us to the second source of foreign accents which assumes that errors in L2 production are caused in part by the inaccurate perception of L2 segments (Flege 1992). For instance, Bever (1981) assumes that segmental speech production and perception develop independently after the critical period in that individuals may learn to discriminate sounds that they cannot still produce accurately. Flege (1995) claims that the production of an L2 phonetic segment will not be more accurate than its perceptual representation and might be less accurate in early stages of development. In other words, the accuracy of L2 speech production is limited by the accuracy of speech perception.

2.1.2 Production of L2 syllable structure

Languages vary in the shape of their syllables. The syllable is "an abstract phonological constituent without clear phonetic correlates" (Zec 2007:161). A syllable's basic role is organizing segments into sequences (Selkirk 1982). Languages of the world permit syllable types with varying degrees of complexity (Levelt and Van de Vijver 2004). Belvins (1995) provides 12 examples of syllable types from the inventories of different languages (Table 2.1).

Language	Syllable Type	Language	Syllable Type
Hua	CV	Klamath	CV(C)(C)
Thargari	CV(C)	Mokilese	(C)V(C)
Cayuvava	(C)V	Totonac	C(C)V(C)(C)
Arabela	C(C)V	Finnish	(C)V(C)(C)
Sedang	C(C)V(C)	Spanish	(C)(C)V(C)
Mazateco	(C)(C)V	Dutch	(C)(C)V(C)(C)

Table 2.1 Different syllable types from the inventories of different languages

Some languages allow simple syllable structures like Hua while others allow more complex structures like Dutch. The CV syllable is the syllable type that all languages have in common, and this type is generally accepted as unmarked with regard to syllable shape (Rice 2007; Zec 2007).[5] The existence of syllables with consonant clusters in a language implies the existence of the CV syllable within that language (Cairns and Feinstein 1982). Also, longer clusters are considered more marked than shorter clusters (Anderson 1987). Moreover, according to Greenberg (1978), consonant clusters appear less frequently in syllable final position. Thus, coda consonant clusters are considered more marked than onset consonant clusters (Anderson 1987). Given that languages differ in syllable structure complexity, L2 learners from language backgrounds that are typologically different from the L2 will face difficulty learning the syllable structure of a new language as will be evident from the research presented below.

[5] This means that it is more natural, more universal, simpler, more common and basic (see Rice 2007 for more markedness terms).

Learning a second language involves adjusting the L2 syllable structure to fit that of the L1 (Broselow 1988). Studies that have investigated the acquisition of an L2 syllable structure, precisely onset and coda consonants, have consistently provided evidence on the prominent role of L1 transfer as well as universals in the construction of an L2 syllable structure mainly in the early phases of L2 acquisition (Young-Scholten and Archibald 2000; Hansen 2004).

Carlisle (1988, 1991) examines the production of English onset consonant clusters of the structure /sC(C)-/ by Spanish speakers of English. Spanish has an epenthesis phonological rule in which /e/ is added at the onset of a large number of words beginning with /sC(C)-/ (e.g. *espia and escuela)*. The study found that the speakers transferred the Spanish rule to English, pronouncing words such as steep and snow and as [estip] and [esno]. Similarly, Broselow (1983) found in a study that L1 transfer played a predominant role in the interlanguage phonology of Iraqi and Egyptian speakers of English. Despite different predictions the two Arabic varieties of the learners' made regarding the position of the vowel inserted in the English syllable structure, speakers from both varieties brought their L1 underlying syllable structures into conformity with surface structure limitations of the L2. Egyptian learners applied their L1 rule of anaptixis and inserted a vowel to the right of the extra syllabic consonant producing an English word such as *flow* as [filo]. Conversely, Iraqi learners used a rule of prothesis and inserted a vowel to the left of the extra syllabic consonant producing *flow* as [iflo].

Other than L1 transfer, a number of studies have also reported that typological universals and markedness contribute to the development of a non-native phonological system (Hansen 2004; Carlisle 2001). The role of universals in the development of an interlanguage phonology originates from the notion that L2 adult learners have access to the parameters and principles of Universal Grammar (UG) with respect to phonology. The most extensively studied universals related to the L2 syllable structure are the CV syllable, the length of the margins and the Sonority Sequencing Principle (SSP). The first two universals relate to the notion of markedness, i.e. naturalness as explained previously. The SSP is also related to markedness and suggests that in all languages of the world, the nucleus is the most sonorant constituent and segments at the margins rise in sonority as they come close to the nucleus (Clements 1990). Hence, according to the SSP, the sequence [obstruent+liquid; e.g. /sl-/ *slime*] would be preferred for an onset of a syllable, while the sequence [liquid+obstruent; e.g. /lt-/ *melt*] would be preffered for a coda of a syllable (Altenberg 2005).

To this end, studies of L2 syllable structure have focused predominantly on an explanation and a description of the type of errors that L2 learners produce. Variation in the production of consonant clusters and the difficulty L2 learners may face when producing them has been mainly investigated from a linguistic point of view by examining the role of L1 transfer or universals. Factors pertaining to the learners social or cognitive status are understudied in this research.

2.2 L2 Speech Perception

The ability to distinguish the acoustic cues underlying the phonemic contrasts of a language starts in infants as young as 1 month of age (Eimas *et al.* 1971). Phonemic contrasts are defined as "segment-sized constellations of phonetic properties that have become linguistically distinctive because they are used systematically to convey differences in word meanings" (Best 1992:1). Languages have both contrastive and non-contrastive segments, and whether a particular phonetic dissimilarity is contrastive or not varies cross-linguistically. For example, Japanese and Korean sound inventories lack the liquid contrast /l/-/r/ which is found in English. In French, vowel nasality is contrastive so words with an oral and a nasal vowel form minimal pairs that differ in meaning (e.g. bas /bæ/ 'low' and banc /bæ̃/ 'bench'). However, nasality is not contrastive for vowels in English (Seidle *et al.* 2009). Even on occasions where two languages have a common phonemic contrast, the phones involved often vary in their articulatory properties. To clarify, Best (1992) notes that both French and English distinguish between /p/ and /b/, voicing of English /b/ may either start concurrent with the release of the closure producing [b] or slightly after the burst and thus /b/ is realized as a short-lag unaspirated [p]. French /b/, on the other hand, is consistently realized as [b] regardless of context. In English /p/, aspiration can either begin after a long release lag and thus the phoneme is realized as [pʰ] or after a shorter lag producing unaspirated [p] in specific contexts. French /p/, however, is consistently realized as the unaspirated short-lag [p]. Hence, English contrasts slightly different phonetically-realized phonemes /b/-/p/ than French does, whereas French does not contrast [p] and [pʰ] in the same way as English does. Another example is the phoneme /r/ which is found in both American English and Spanish but is realized as a liquid approximant [ɹ] in the former and as an alveolar tap [ɾ] or trill [r] in the latter (Best 1992).

By listening to natural language input, infants are not just responsive to speech sounds but are also able to categorize phonemes (categorical perception) in a manner similar to how adults perceive these phonemes (Eimas *et al.* 1971; Kuhl *et al.* 1992; Jusczyk *et al.* 1993; Kuhl 2000). Infants' initial capacity to detect differences between the phonetic units is found to be innate and universal (Eimas *et al.* 1971; Kuhl 2000). Evidence for this comes from findings on infants' ability to discriminate phonetic units of native and non-native speech. Speech experience then tunes infants' perceptual system to the phonological properties of the ambient language before lexical learning takes place, that is, before children understand or produce words (Kuhl *et al.* 1997).

Research shows that by 12 months of age, infants' ability to discriminate non-native phonemic contrasts declines considerably (Werker and Tees 1984; Best and McRoberts 2003; Kuhl 2000; Kuhl *et al.* 2006). As to why speech perception changes from the language-general to language-specific, Eimas and Corbit (1973) hypothesize that underlying the perception of speech are phonetic feature detectors that are differentially sensitive to phonetic units in speech and thus can be maintained or lost due to experience. Detectors triggered by speech input are maintained while those not triggered by speech atrophy (Eimas 1975). Research on brain imaging alternatively shows that language acquisition involves "neural commitment" that is associated with the well-established issue of a "critical" period (Kuhl 2007:71). The brain's neural networks code native language speech early in development and ultimately make it hard to learn patterns of a new language (ibid). The change in infants' perceptual abilities due to experience is well demonstrated by a study examining the perception of Hindi dental/retroflex initial stop consonants by English infants, children and adults. Results of cross-sectional and longitudinal experiments provided evidence that both Hindi and English 6-month-old infants could successfully discriminate this consonant contrast (Werker and Tees 1999). However, children over 4 years of age and adults failed to discriminate the Hindi contrast, which is not phonemically contrastive in English. Not only does infants' capacity to detect non-native contrasts decline by age, but their native language phonemic perception also becomes stronger as they grow up (Kuhl 2000; Kuhl *et al.* 2005). Another well-documented and widely cited example of adults' difficulty in perceiving non-native contrasts is that of Japanese listeners' discrimination and identification of the English /ɹ̠/-/l/ contrast. Several studies using synthetic or natural speech material and a variety of tasks (e.g. identification, discrimination) found that the performance of native Japanese listeners for whom the contrast is not phonemic is often significantly poorer than that of English listeners (e.g. Best and Strange 1992; Akahane-Yamada 1995; Iverson and Kuhl 1996).

Despite evidence on adults' difficulty in discriminating foreign language phonemic contrasts, some research has provided evidence that this ability is not completely lost. Some studies have suggested that adults' performance on perceiving non-native phonemes can be increased by the implementation of techniques that minimize the influence of memory or the use of extensive training trials (Werker 1995). For instance, research has provided evidence that Japanese learners of English are able to successfully perceive and produce the English /ɹ/ and /l/ contrast under laboratory training conditions and using a 'high-variability' training method[6] (e.g. Barriuso and Hayes-Harb 2018; Shinohara and Iverson 2018). Other research has demonstrated that not all phonemic contrasts are equally difficult for L2 learners to perceive, and adults are sometimes able to distinguish non-native contrasts without any training (Best 1995, 1993). One example comes from a study by Best *et al.* (1988) who tested the ability of English infants and adults to discriminate the apical/lateral click contrast in the Zulu language. The study found that English participants of all ages easily discriminated the contrast contrary to the claims of the maintenance/loss model. This finding was explained in light of the Perceptual Assimilation Model developed by Best and colleagues (Best *et al.* 1988; Best 1995, 1993). PAM claims that non-native sounds are perceptually assimilated to native sounds based on how listeners detect similarities or dissimilarities in the properties of the articulators (constriction degree, active articulators, constriction locations, phasing) (Best *et al.* 2001). That is, the non-native sound will be heard as a good or a poor exemplar of a native phoneme, or as different from any native phoneme or, infrequently, as a non-speech sound (Best and Tyler 2006). American speakers' successful discrimination of the Zulu click contrasts was found to support the prediction that these contrasts were perceived as non-speech sounds.

Also, the Speech Learning Model developed by Flege (1995) argues that the mechanisms that children use when learning the L1 remain intact across the life span and continue to be available for use in L2 learning. This, however, does not mean that children and adults will eventually attain the same proficiency in the L2. According to SLM, most L2 learners continue to differ from young learners because they keep using their L1 which interfers with their L2 learning.

[6] This approach exposes participants to the whole set of variable stimuli within each contrasting category that learners can encounter in real input (Bradlow 2007).

In addition, as L1 categories become more established with age, difficulties in perceiving and producing the L2 arise.

From research on speech perception and production, we know that the most investigated factors in L2 speech learning are age and L1 influence. Factors such as education and literacy have been isolated in most studies of L2 perception and production. Theoretical models that have developed in this area such as PAM, SLM, the Native Language Magnet Model (NLM; Kuhl 1991; Kuhl and Iverson 1995), the L2 Version of the Linguistic Perception Model (L2LP; Escudero and Boersma 2004; Escudero 2006) and the Automatic Selective Perception (ASP; Strange and Shafer 2008) have two main emphases: (1) predicting relative challenges in perceiving and producing L2 segments and contrasts, or (2) considering the exact processing mechanisms in L2 phonetic perception (See Jing *et al.* 2019 for a review). Hence, testing these models requires an evaluation of the similarity between the phonetic aspects of the L1 and the L2 sounds.

2.3 Studies on L2 Arabic Speech Learning

Very few cross-linguistic studies exist that investigate the perception and production of L2 Arabic consonants and vowels by adults. Shehata (2015) reviewed an early study by Asfoor (1982) who investigates the production of Arabic segments by native adult English speakers and the role of the learners' dialect in their production patterns. 24 instructors of Arabic at the Defense Institute in Monterey were asked to evaluate the 10 most difficult Arabic sounds produced by their students in various word positions: initial, medial and final. 34 students were recorded producing these sounds prior to and after a six-week training course. Findings revealed that stops were the most difficult for English learners of Arabic to produce (the exact sounds were not mentioned in this review). Also, training was found to promote the learners' pronunciation of challenging sounds including /ðˤ, χ, ʁ, ʕ/).

Along similar lines, Alwabari (2013) examins the production of the pharyngeal consonants /ʕ, ħ/ and their plain counterparts /h, ʔ/ as well as the pharyngealized /tˤ, sˤ, dˤ, ðˤ/ and their plain counterparts /t, s, d, ð/ by adults learning Arabic who varied in their proficiency level. The study recruited an experimental group consisting of English native speakers who had been exposed to Arabic via Arabic classes, and a control group consisting of English native speakers who had not had regular exposure to Arabic or any Semitic language. A rapid shadowing paradigm was used to test speakers' accuracy in producing the target sounds, in which speakers heard an audio stimulus containing the target sound and repeated what they heard immediately. The study found that English native speakers with higher proficiency in Arabic were more accurate in producing Arabic pharyngeals and pharyngealized consonants compared to the control group. The results also revealed that subjects obtained significantly higher accuracy scores with regard to the production of the non-pharyngealized consonants compared to the pharyngealized ones regardless of their Arabic proficiency level. Also, the subjects' mean scores on the pharyngeals' production were significantly higher than their mean scores on the pharyngealized productions.

The aim of the aformentioned studies was to examine non-native adult speakers' production accuracy per se. A few other studies examined non-native perception of Arabic segments. Alosh (1987) examines the strategies used by American English learners of Arabic to perceive Arabic fricative pharyngealized consonants and compared them to those used by a control group of Arabic speakers. The researcher predicted that non-native learners of Arabic perceive pharyngealized consonants using cues from vowels following them while Arabic speakers rely primarily on cues from the consonants themselves to perceive pharyngealization. The study also sought to examine the effect of learners' proficiency level in their perception accuracy and the relationship between their perception and production of the consonants examined. The study used synthetic speech material to test for the effect of vowel quality on learners' perception. 36 American English participants who were divided into three groups based on the extent of Arabic experience they had (this was based on the total number of credit hours of Arabic classes) were recruited for the study. In addition, 10 Arabic speakers were recruited as a control group. In a laboratory setting, the subjects had to respond to a forced-choice task in which they were asked to listen to each token and then judge whether it corresponded to any of the choices they had on a printed sheet of paper. The choices included all tokens in the test (e.g. su:, si:, sa:, $s^ʕ$a:, $s^ʕ$i:, su:). The same syllables used in the perception experiment were then used in the production experiment in which subjects were asked to produce the syllables loudly on a microphone. The findings revealed that the NSs of Arabic outperformed the learners in the perception task and were more sensitive to cues of pharyngealization either on the consonants themselves or the vowels than the learners. Furthermore, learners were more successful in perceiving pharyngealization when the consonant was followed by a pharyngealized vowel rather than when the segment was plain. Another finding this study revealed concerns the effect of the learners' proficiency level. Perception of pharyngealization was found to be significantly higher as the proficiency of the learners advanced. As for production, the study found non-native learners of Arabic did not differ in their production of the pharyngealized consonants based on proficiency level. It also revealed no significant relationship between production and perception scores. The author speculates that this does not provide support to the notion that perception is dependent on production.

More recently, Al Mahmoud (2013) examines discrimination of Arabic consonant contrasts by American English-speaking learners of Arabic in light of PAM. 22 students enrolled in an Arabic programme carried out an AXB forced-choice task in which they were asked to indicate whether the first or third test item (A or B) was the same as the second (X) or not on a sheet of paper. The results revealed that Arabic learners were able to discriminate contrastive pairs that contained consonants that were assimilable (equivalent) to two separate L1 categories (e.g. /t/-/d/ and /θ/-/ð/). This finding supports PAMs' prediction that discrimination will be excellent when each non-native sound is assimilated to a different native category. In the contrary, contrastive sounds such as (/x/-/ɣ/, /ħ/-/h/, /x/-/ħ/) were poorly discriminated by the learners.[7]

The previous studies draw our attention to two main conclusions: (1) Arabic proficiency level of the learners is a primary determinant of the L2 production and perception and (2) success in perceiving non-native contrasts depends on their similarity to or difference from native language phonemes. The studies varied in their methodologies and the variables examined and their conclusions are applicable to the setting and methodologies these studies adopted. The previous studies examined non-native Arabic production and perception of foreign language learners, in these cases university students enrolled in Arabic courses in a predominantly English-speaking environment. However, we know little about how non-native learners of Arabic perceive or produce Arabic sounds when they are exposed to Arabic in a naturalistic setting and receive no formal training on the language (See more about L2 learning setting in 2.4.3). This will be explored in the present study.

[7] Other than an effect of the learners' native language, the author also attributed poor discrimination of some contrasts such as /x/-/ɣ/ and /x/-/ħ/ to the orthographic similarity between the two consonants in a pair (something that has been overlooked by PAM).

2.4 Factors Affecting L2 Speech Learning

A vast amount of research has discussed the factors that might influence the acquisition of L2 phonology. In addition to age and L1 influence, which have already been introduced, in the following paragraphs, I will briefly discuss the factors that are more relevant to the current study.

2.4.1 Length of residence

In naturalistic settings, one way that studies have used to statistically measure the effect of input on L2 acquisition is to account for the period the learner has spent living in the native-language community. Length of residence (LoR) is a simple, continuous variable that has been frequently studied in SLA research as an index for ultimate attainment in L2 phonology (Flege 2009; Moyer 2009). Flege (2009) postulates that if input to L2 learners matters, then LoR should be correlated with measures of L2 speech attainment. Likewise, McAllister (2001) states that there is a general assumption that LoR correlates positively with the amount of input an L2 learner has acquired and that the more L2 input one receives the better the opportunities for the L2 learner to master the L2. However, reports on the significance of LoR effects in the existing literature are inconsistent (Piske *et al.* 2001; Dekeyser and Larson Hall 2005). Several studies examining the role foreign accents and production accuracy of L2 phonetic segments have supported the conclusion that the influence of LoR on L2 performance is insignificant (Moyer 1999). Oyama (1976) found no effect of LoR on the L2 phonology of Italian English speakers in the United States when the effect of their AoA was controlled for. Similarly, Flege (1988) reported no difference in foreign accents between two groups of Chinese adult immigrants to the United States based on their LoR, which varied between 1.1 and 5.5 years. In another study by Flege (1993), Chinese adults with average LoR in the United States between 1.2 and 5.1 years did not differ significantly in their English production of sounds.

As to why LoR has not been found to affect L2 speech learning in previous studies, Flege and Liu (2001: 531) provide possible explanations. First, the amount of L2 input is not a critical determinant of L2 speech learning. Second, as most participants in previous studies were late adult L2 learners, L2 performance is constrained by a sensitive or critical period. Third, the nature and amount of L2 input affect L2 performance but "LoR provides a good index of L2 input only for certain individuals". In support for the third reason, Oyama (1976) observes that for some individuals, social and economic factors can inhibit L2 acquisition even if they are capable of learning the L2. Also, LoR might provide a good index of L2 input to children immigrants more than adult immigrants. L2 performance of immigrant children increases substantially over time (Cummins 1991). This is because immigrant children have regular contact with NSs as they get enrolled in schools that are mostly populated by native speaker children (Flege and Liu 2001).

Another explanation for the lack of an LoR effect in previous research is provided by Moyer (2009) who argues that the influence of LoR has been under-explained because as an independent measure, it does not reveal much about the quality of the target language input. Thus, one reason for the small or absent effect of LoR is that residence in a country does not guarantee authentic input or interaction. Evidence for this comes from a study by Flege *et al.* (1997) who examin the role of experience in the production and perception accuracy of English vowels by 20 immigrant speakers to the United States from different L1 backgrounds. Acoustic measurements of production data revealed that LoR played a significant role in the subjects' accuracy levels. Nevertheless, the authors observed that the performance of experienced learners was not entirely native-like and that phonetic learning was not manifest in all the data examined. The authors noted that factors other than experience (e.g. L1 influence, psycho-social factors and speaking the L2 with other non-native speakers) could have influenced the subjects' attainment levels. Another study that ties to this dispute is a comparison of the performance of 60 Chinese student- and non-student-immigrants to the US on grammaticality judgment, listening comprehension and phonemic recognition tests (Flege and Liu 2001). The subjects' L2 proficiency level was found to increase with LoR only for the students. The authors speculated that students were exposed to more authentic input through their interaction with teachers and peers, and therefore, the impact of their LoR was significant compared to the non-students. From this, it is obvious that the length one spends living in a country does not necessarily lead to better attainment and that other factors such as practice and quality of input play a critical role (Moyer 2009).

A few other studies found that LoR plays a positive role in L2 performance. For example, a study by Flege (1988) revealed that Spanish adults who had lived in the United States for an average of 14.3 years had significantly better pronunciation of English sentences than individuals with an average LoR of 0.7 years. Flege *et al.* (1997) found in another study that LoR played a modest effect on the perception of English vowels by adult NNSs of English who varied in their LoR between 0.7 and 7.3 years. They also reported a significant effect of LoR on the accurate production of one English vowel among others examined.

With the recognition that LoR alone does not help guarantee any particular quantity or quality of L2 input, researchers have started examining more comprehensive aspects of L2 input. One way to measure the significance of input beyond LoR is to look at opportunities for language contact and use (Moyer 2009). Studies confirm that a substantial amount of time spent communicating with native speakers is highly beneficial for accent (Flege *et al.* 1995; Moyer 2013). Also, living with a family is more likely to enhance L2 fluency because it ensures a continuous setting for interactive language use (Moyer 2009). Flege *et al.* (1999) note in their study that despite the usefulness of extended exposure to a foreign language in a classroom setting, ultimate attainment in learning a foreign language is also dependent upon a considerable amount of native-speaker input. In instructed settings, exposure to instruction is useful in improving one's accent, particularly when tasks are targeted to improve learners' segmental and suprasegmental accuracy (Moyer 2009). However, Young-Scholten (1995) points out that input provided in foreign classrooms can have negative consequences if that L2 input comes from teachers or peers who are foreign-accented. Immigrants can also be exposed to a substantial amount of foreign-accented input when they spend most of their time with other L2 non-native speakers (Piske *et al.* 2001).

2.4.2 Education and literacy

This section will provide an orientation to research conducted on adult L2 learners who are non- or low-literate due to little or no formal education and will focus specifically on the relationship between literacy and second language achievement.

None-literate and low-educated second language learners have been part of SLA research since the 1970s (see Craats *et al.* 2006 for a review). This research included uninstructed immigrant adults to the USA, Europe and Australia, who had no other linguistic competence than that of their native language and who had varying levels of formal education (mainly primary). The focus of these studies was on the extent to which adult L2 learners could acquire language like children, based merely on aural input, and therefore used naturalistic L2 adult learners as in child L1 acquisition. Language acquisition (mainly of morphosyntax) of adult immigrants was tested based on production data. For instance, the ZISA (Zweitspracherwerb Italienischer

(Portugiesischer) und Spanischer Arbeiter) project in Germany reports on the findings of a cross-sectional and longitudinal study (Clahsen *et al.* 1991). The study investigates the acquisition of syntax by 45 Italian, Spanish and Portugese immigrants. The learner's performance was examined for 2 and a half years. Another longitudinal study is Cancino *et al.*'s (1978) study in the USA. This study examines the acquisition of English negatives and interrogatives by 2 children and 2 adolecent Spanish immigrants for a period of 10 months. Moreover, the ESF project (European Science Foundation), which combined six researchers examined the acquisition of morphosyntax by 40 immigrant adults from 6 L1 backgrounds including Punjabi, Moroccan Arabic, Turkish, Spanish, Finnish and Italian for a period of six years (See Perdu 1984 and Perdue 1993). The L2s examined were five European languages including English, Dutch, French German and Swedish. The project adopted longitudinal methodology, which extended for 10 months. Other studies used cross-sectional methodologies only, such as the Lexlern Project and the Heidelberger Pidgin Projekt in Germany (Klein and Dittmar 1979)

One major finding the aformentioned studies have contributed is that L2 learners follow a common route in their L2 development of morphosyntax regardless of their age, educational background, L1 or input. Despite the significance of these studies, most have not addressed so many other issues that have been ignored in the wider SLA research including the impact of literacy or education on L2 acquisition (but see Young-Scholten and Storm 2006 and Kurvers *et al.* 2006 for research on the impact of literacy and metalinguistic awareness on learning reading by low educated adults).

One early study of naturalistic acquisition that has looked at non-linguistic factors including educational level is that of the Heidelberger Pidgin Project (Klein and Dittmar 1979). The study was concerned with examining the oral skills of a group of uninstructed adult immigrants in Germany. The study recruited 24 Italian- and 24-Spanish speaking foreign worker immigrants. The subjects who were all above the age of 18 years, were divided based on their LoR in Germany into four groups (up to 2 years, 2-4 years, 4-6 years and above 6 years). Oral data was collected via direct conversation. The subjects were placed into four proficiency groups based on their production accuracy of 100 consecutive utterances. In addition, a control group was recruited for comparison reasons. Analysis of the learner' L2 morphosyntax revealed that their developmental progress followed the same path as other learners reported in previous studies. However, the learners' proficiency levels varied due to the influence of other factors.

For example, variation in performance was accounted for by factors including the type of job, frequency of contact with native speakers, location of residence and years of formal education. The study also found that those who belonged to the lowest proficiency level had not finished primary schooling. However, the relationship between schooling and proficiency level for those who had a greater amount of formal schooling (beyond primary school) was not linear.

SLA studies that have become interested in investigating the relationship between literacy and L2 acquisition or performance are informed by research on the effect of L1 literacy on L1 language processing. Research in the field of native-language cognitive processing has demonstrated that non-literate adults have considerable difficulty solving oral language tasks in their native language that require a focus on the phonemes, syllables or words of the language (Morais *et al.* 1979). For example, Morais *et al.* (1979) and Morais *et al.* (1986) examine the oral language processing abilities of a group of otherwise healthy non-literate Portuguese speaking adults in Brazil. They found that they performed poorly in tasks that required segmental analysis, unlike literate individuals. One typical oral task that was easy for literates but quite impossible for non-literates to solve asked them to delete or add a consonant at the beginning of a word. Differences in performance also appeared in phonological fluency tasks which required a participant to list all the words they could recall that start with ‘p’, ‘t’ or ‘s’. The researchers argued that this difference in performance is due to differences in the mastery of an alphabetic script, which involves the formation of grapheme-phoneme correspondence. Illiterate adults, unlike literate adults, lack knowledge of the alphabetic script and therefore find it hard to perform oral tasks that involve the construct ‘phoneme’. Another study examined two groups of Chinese with comparable age, education and social status but one was literate in Chinese characters only while the other was literate in both the Chinese characters and Chinese alphabetic script (Read *et al.* 1986). The groups were asked to perform a similar oral task to the one in the previous studies. Results revealed that those literate in the alphabetic script performed significantly better than the other group. The researchers conclud that the ability to read and write alphabetically enhances the ability to segment oral language.

Subsequent studies were carried out in Brazil, Portugal, and Spain to compare the performance of literate and non-literate adults. For instance, Adrian *et al.* (1995) conducted several oral tasks for a group of 15 non-literate Spanish adults and compared their performance to two other groups: “poor readers” and “readers”. One aim of the study was to test the groups’ phonological sensitivity using a same-different classification task in which subjects were presented with syllables differing in the first consonant such as /me-me/ and /ta-sa/. Other tasks involved phoneme and rhyme detection, phoneme and syllable deletion and phoneme, word and syllable reversal. Results revealed that all learners performed perfectly well in the phonological sensitivity task. However, for the other tasks, illiterate adults obtained very poor scores compared to the other two groups. The authors explained that the high scores of illiterate adults in the discrimination task suggest that phonological sensitivity differs from phonemic awareness. They argue that phonological sensitivity is a universal competence because it is "a basic component of speech processing" (p.331). However, the other tasks in which the illiterate adults performed poorly require conscious phonemic awareness and thus require knowledge of reading and writing the alphabetic script.

The previous research provides evidence that confirms that knowledge of the alphabetic script provides essential cognitive tools that aid the processing of oral language (Tarone and Bigelow 2005). In support of this, Young-Scholten (2013: 446) states that knowledge of the alphabetic script is "a metalinguistic activity that requires a sophisticated level of awareness". Tarone (2010) considers the implications of this conclusion in SLA research. She argues that if L1 literacy significantly affects oral language processing, then it should also affect the way L2 input is processed during the acquisition of morphosyntax. As there are cumulative numbers of non- and low-literate adults migrating to highly literate countries, such as Oman, understanding their cognitive limitations in terms of language processing is critical. Because SLA research has relied mostly on highly literate adults, normally university students, in drawing conclusions on how L2 input is processed, theories that have developed in this realm may not apply to non-literate L2 learners (Tarone 2010; Tarone and Bigelow 2005).

The focus of SLA research which has considered the role of literacy has been on the effect of literacy on the development of metalinguistic awareness especially morphosyntax (e.g. Kurvers *et al.* 2002; Tarone *et al.* 2009; Janco *et al.* 2019) and phonology (e.g. Young-Scholten and Strom 2006; Young-Scholten and Naeb 2010). One major finding from this research is that age is not a barrier in the acquisition of morphosyntax or learning to read for the first time, but that literacy of the language learner plays a role in this matter. However, research that investigates the effect of literacy and education on L2 speech learning or phonological acquisition is scarce. One recent study by Saleh Al Azmi (2019) examines the role of L1 literacy on the acquisition of a complex English syllable structure. The study recruited 60 Najdi Arabic speakers, 20 of which were non-literate in Arabic and all 60 were non-literate in English. The learners were given English lessons by exposing them to English words containing onset and coda clusters. The learners were divided into three groups during the lessons: (1) Arabic non-literates were exposed to aural input only, (2) 20 Arabic-literates were exposed to aural input only, and (3) 20 Arabic literates were exposed to both aural and orthographic input (See next section for more details on orthographic input). From a post test, the study found that non-literate Arabic speakers relied on consonant deletion to produce English consonant clusters, but literate Arabic learners who were exposed to aural input relied more on vowel epenthesis in cluster production. The author explains this difference in phonological errors between the two groups by comparing the non-literate group to children who rely on consonant omission as a strategy in their production of L1 consonant clusters. Hence, both the literate and non-literate group were only different in the strategy they employed in production of the target syllable.

2.4.3 Input

From early accounts on the critical period, we now acknowledge that the input which learners are exposed to in the target-language community is one of the most crucial factors that can influence the rate and end-state of L2 learners' attainment level (Best and Tyler 2006; Piske and Young-Scholten 2009). For example, a central motivation for the study of input in SLA research has been the belief that the target language input in itself is insufficient for successful language acquisition but rather what is necessary is comprehension of that input by the learner (Pica *et al.* 1986; Pica *et al.* 1987). Hence, the Interaction Hypothesis proposes two major claims with regard to the role of input in SLA: (1) comprehensible input is indispensable for L2 acquisition, (2) adjustments to the input that occurs in the process of negotiating meaning during communication facilitate comprehension of the input to the L2 learner (Long 1981; Long 1985). Another example for the importance of input, which is more relevant to the present thesis, comes from speech learning research. L2 speech perception research has found that the role of age in L2 perceptual learning takes place mainly through interaction with the amount of exposure in the L2, including LoR, quantity and quality of input and relative usage of the L1 and L2 (Best and Tyler 2006).

Input is available to first language learners through interaction with adults and other speakers of the language. For second language learners, information about the language to be learned is available in different forms that can be generalized into three areas: (1) naturalistic input, (2) formal instruction, (3) a combination of 1 and 2 (Pica 1983) (See Section 2.4.1 for some differences between the two settings). Additionally, L2 learners are often exposed to spoken or orthographic input (or both) (Bassetti 2017). Orthographic input consists of the sounds and words in a written form (Bassetti *et al.* 2015). L2 orthography affects L2 speech learning in different ways. It often facilitates perception, production and word learning (e.g. Showalter and Hayes-Harb 2013). It can, on the other hand, lead to non-native pronunciation (e.g. Bassetti 2007; Young-Scholten and Langer 2015), or it can play no role at all (Bassetti *et al.* 2015).

Other than the amount or form of the target language input, the communicative and phonetic characteristics of the language input seem to affect learners' developmental level and L2 comprehensibility (Best and Tyler 2006; Gass and Mackey 2007). With regard to communicative features, the Interaction approach has shown that L2 learners are more exposed to modified input (Gass and Mackey 2007; Maleki and Pazhakh 2012). Modified input can be generated in different forms, such as repetitions, paraphrasing and reduction of utterance length or complexity. Thus, this research has shown that adjustments made by speakers in terms of simplifications or elaborations made the language more comprehensible and easier for the learners to access (Gass 2008; Gass and Mackey 2007). With regard to phonetic properties, infant-directed speech (IDS) contains hyper-articulated vowels (e.g. Kuhl *et al.* 1997) and consonants (e.g. Cristia 2010) that are assumed to serve a didactic role, which can help the infant learner to acquire the native language (Chapter 3). Foreigner-directed speech (FDS), similarly, exhibits vowel hyper-articulation but not the emotional prosody that IDS serves (Knoll and Uther 2004; Uther *et al.* 2007). Hence, systamatic adaptations in speech directed to language learners are essential in demonstrating how the linguistic environment interacts with the learners' linguistic and developmental state (Best and Tyler 2006). Despite the importance of this research, we still know little about the phonetic properties of FDS. The present study will primarily explore the acoustic properties of input directed to an understudied group of L2 learners.

2.5 Conclusion

This Chapter presented research on the production and perception of speech by L2 learners. It highlighted the difficulties that L2 learners face in this process. It also dicussed the factors that play a role in SLA including LoR, literacy and education as well as input. Now that we know how L2 learners percieve and produce phonological and phonetic input, the following Chapter will be dedicated to discussions on the properties of phonological and acoustic input to different audiences including L2 learners.

Chapter 3. Linguistic Adaptations and NS-NNS Interaction

3.0 Introduction

This chapter will review the literature on infant-directed speech (IDS), clear speech and foreigner-directed speech (FDS) and will present the linguistic modifications reported in each speech code. It will present general linguistic features of each code and then will elaborate specifically on the phonological, phonetic and prosodic features to provide a background to the current study. Later, the Chapter will present the most relevant theoretical models that can be used to interpret the dynamics of NS-NNS communication, including the Accommodation Communication Theory and the Hyper and Hypo theory.

3.1 Infant-Directed Speech

Linguistic studies on speech addressed to children started as early as in the mid twentieth century, when Ferguson (1964, 1971, 1975) conducted some investigations of special registers (e.g. 'baby talk' and 'foreigner talk') which are regarded by community groups as 'imitations' or 'simplified' versions of the language used with people who are considered as unable to comprehend the target speech of the community. Ferguson (1964:103) defines 'baby talk' (also referred to as 'parentese', 'motherese' and more recently 'IDS') as "any special form of a language which is regarded by a speech community as being primarily appropriate for talking to young children and which is not generally regarded as the normal adult use of language". Other registers are also observed in other circumstances, such as when addressing a family pet (e.g. Mitchel 2001, 2004), or when conversing with others who lack power but who trigger affection such as the sick (e.g. Levin *et al.* 1984) or when conversing with the elderly (e.g. Ryan *et al.* 1994). According to Ferguson (1975:1), a fundamental theoretical justification for the study of simplified speech registers (e.g. baby talk, foreigner talk, talking to deaf) is the implications they might have for such concepts as 'basic', 'simple' or 'deep', which are central notions in theories of language, in addition to understanding language development in the case of baby talk and pidginization processes in the case of foreigner talk.

IDS is described to contain short and enunciated utterances and use expressions that concern events or things that are present rather than past or absent (Gleitman *et al.* 1984; Soderstrom 2007). IDS is also characterized with hyperarticulation, slow rate, elevated pitch and distinct prosody (Eaves *et al.* 2016; see Sections 2.1.2 and 2.1.3). Infants show a preference for IDS and are able to distinguish IDS from typical adult-directed speech (ADS) (Pegg *et al.* 1992). The following sections will expand more on IDS, particularly its phonological, phonetic and prosodic characteristics that have been investigated in the literature.

3.1.1 Phonological adaptations in IDS

Some early research on baby talk (e.g. Ferguson 1964; Kellar 1964; Bynon 1968) provid descriptive accounts of speech directed to infants by relying on informant or parental reports. The main aim of these studies was to compile lists of words from a language or different languages and note similarities between them by comparing them cross-linguistically. Ferguson (1977) distinguishes three different functions of baby talk including simplification, clarification, and expressiveness. Simplification is associated with various segmental modifications, whereas the clarification and expressiveness functions are associated with prosodic features (Cruttenden 1994). In one account, Ferguson (1964) investigates baby talk among six languages (Syrian Arabic, Marathi, Coman-che, Gilyak, American English, and Spanish). He points out that phonologically, baby talk is characterized by simplification of consonant clusters in all languages examined except Arabic (though he noted that it could also occur here). However, Ferguson does not provide information about the strategy of consonant cluster simplification (either deletion or epenthesis) or examples for it. Furthermore, Ferguson provides examples of how parents from all languages attested replace some consonants with others; for example, parents realized the sound /r/ as [l, w, j, d, or t] as in the English word "*wabbit* for *rabbit"* (p.105). Interchange among some consonants (e.g. affricates, sibilants and stops) also occurrs in some languages. For instance, parents realized /ʃ/ as [s] as in English "*soos* for *shoes"* and /s/ as [tʃ] as in Spanish "*becho for beso"* (p.105). Ferguson reports other phonological features including loss of unstressed syllables and distant nasal assimilation. Other studies also report consonant cluster modification in baby talk. According to Cruttenden (1994), consonant clusters reduce to one consonant in baby talk; for example, English clusters of /s/+C and C+approximant are often reduced to the consonant alone (e.g. drink=[dɪŋkɪ]). Rūķe-Draviņa (1977) reported that consonant clusters rarely occur in Latvian baby talk. Furthermore, the consonant-vowel structure has been found to be preferred in baby talk. A combination of consonant harmony and reduplication result in structures like [dʒi dʒi] for English *horse and* [ɣu ɣu] for [aɣu] ("milk") in Berber (Bynon 1968). According to Cruttenden (1994), the consonantal and syllabic simplifications that occur in baby talk indicate that parents are showing solidarity with their children by producing speech that is more familiar to them.

3.1.2 Phonetic enhancement in IDS

A considerable amount of research has investigated whether the input available to infants predisposes them to learn linguistic categories. Much of this work has studied the acoustic features of IDS. Given the new discoveries on the importance of ambient input to infants' perceptual development (Chapter 2), research has taken into account how IDS might be an important source for the development of their perceptual abilities.

3.1.2.1 Studies on consonants

One line of research has reported that IDS involves clearer or exaggerated cues to consonantal contrasts, like that of voice onset time (VOT). This research draws from theories of speech perception, especially findings on infants' ability at categorizing speech sounds from the environment. For instance, Eimas *et al.* (1971) demonstrates in his study that young infants, like adults, exhibit a categorical perception of the VOT feature of phonemes with relatively little exposure. Subsequent studies considered whether maternal speech exhibits exaggeration of the VOT contrast by comparing it with adult-directed speech (ADS), with conflicting results. While some studies provided positive evidence for consonantal enhancement of the voiced/voiceless contrast in IDS (e.g. Malsheen 1980; Sundberg 2001; Cristia, 2010; Burnham 2013), others found no difference and sometimes reduction (e.g. Baran *et al.* 1977; Sundberg and Lacerda 1999; Synnestvedt 2010). Malsheen (1980) investigates the speech of six English mothers as they engaged in natural interactions with three age groups of children in a longitudinal study. She specifically examined VOT values in voiced and voiceless prevocalic stops in the speech directed to 6-to- 8-month-olds, 15-to-16-month-olds, and 2-to-5-year-olds, and adults. Her results revealed a significant difference in VOT values between IDS and ADS in the speech directed to the 15-16-month-olds with mean VOT in voiceless stops produced in IDS showing higher values than when produced in ADS. Longer VOT values for voicless stops led to a greater difference between voiced and voicless stops in IDS compared to ADS. However, no significant difference was found for the other two groups.

Sundberg and Lacerda (1999) also investigate VOT characteristics of voiced and voiceless stops as evident in the speech of six Swedish mothers once interacting with their 3-month-old infants and another with an adult in a laboratory setting. All instances of prevocalic initial and medial voiced and voiceless stops in stressed and unstressed lexical syllables were analyzed. Unlike Malsheen's (1980) findings, the authors in this study found that VOT values in IDS were shorter compared to those in ADS. Also, VOT mean values of especially voiceless stops were found to be significantly longer in stressed syllables in ADS compared to IDS. The authors explain the reduction in the phonetic specification of stops in IDS in terms of a functionalist mother-infant phonetic interaction model (MIPHI; Sundberg 1998) and in light of the Hyper and Hypo Theory (Lindblom 1990, 1992; see Section 2.3.3 for more details on this theory). They claim that since language is acquired via an interaction between the ambient input and physiological constraints along with other factors such as attention and affection, infants' linguistic development determines the kind of exaggerations in IDS. Prosodic aspects of IDS are over specified at early stages of linguistic development and decrease in specification as the child grows older whereas consonantal properties increase in specification as the child benefits from more ambient input. Sundberg's and Lacerda's findings provide support for the MIPHI model which predicts that consonants unlike vocalic segments are underspecified in the first months of life in IDS and will increasingly become more specified like in ADS at six months of age.

However, a study by Englund (2005) found counterevidence to the predictions of the MIPHI. The author investigates the same linguistic phenomenon longitudinally in the speech of six Norwegian mothers as they interacted with their infants from birth until six months of age. The recordings were carried out at the mothers' home on 10 occasions during nappy change time to ensure naturalistic interactions. After recording IDS, the author interacted with the mother to elicit ADS. Contrary to the findings of Sundberg and Lacerda's, this study found an overall increase in VOT values for voiced and voiceless stops (except for /p/ which was neither enhanced nor reduced) in IDS compared to ADS. Infants' age or linguistic development did not seem to play any role in this phenomenon, as overspecification of stops was evident over the first six months of age. Another study that provid counter evidence to the MIPH is that of Englund and Behne (2006), who examine whether the obstruent /s/ would be underspecified in IDS over the first six months of infants' age based on the MIPHI predictions. After extracting all occurrences of /s/ from data collected from a previous study (Englund 2005), fricative duration was measured and revealed an overspecification of /s/ in IDS during infants' first six months of life compared to ADS, in contrast to the predictions of MIPHI. In other words, /s/ in IDS was constantly longer in duration over the infant's first six months of life compared to that of ADS.

Cristia (2010) attributes the mixed results in previous studies to some theoretical and methodological problems that are worth mentioning in detail here. Previous IDS studies differ in the way they interpret phonetic enhancement. Some studies regard enhancement as an exaggeration (overspecification) along the temporal dimension (e.g. Sundberg and Lacerda, 1999; Sundberg 2001; Englund 2005; Englund and Behne 2006). This could be due to the assumption that temporal exaggeration draws listeners' attention to the emphasized sounds (Cristia 2010). However, this interpretation does not clearly take into account whether the temporal properties considered are perceptually relevant to the category under investigation or how these properties contrast this category from other categories in the language inventory. For example, lengthening of /s/ in Englund (2006) is considered an enhancement regardless of the fact that length does not play a role in distinguishing /s/ from other categories or even in defining /s/ itself as a category. Another interpretation of enhancement emerges from clear speech research (for more elaboration on clear speech see Section 3.2). In this line of research, interpretation of phonetic enhancement relies largely on the sound inventory and arises along perceptually relevant scope, based on the assumption that sounds become more distinct from categories that are neighbouring in phonetic space. As evidence for this, Kang and Guion (2008) report that talkers enhance consonantal contrasts in clear speech along an acoustic dimension that they themselves depend on to discriminate the contrastive sounds. Based on this study, while older speakers rely on the enhancement of VOT to discriminate Korean stops, younger speakers rely on the enhancement of mean fundamental frequency ($f0$) to discriminate the same stops in clear speech. This is because the acoustic correlates of Korean stops had undergone a change in speakers born after 1965.

Variation in the previous results regarding phonetic enhancement could also be due to variation in the linguistic ability of the children examined, rather than just their chronological age. One hypothesis is proposed in Malsheen's study; phonetic enhancement was evident in mothers whose children were at the first-word stage at the time of testing while enhancement was variable in speech to prelinguistic infants and children at the phrase stage. Another hypothesis relates to the nature of the phonological category examined; for instance, Sundberg (1998) suggests that consonantal categories may be underspecified in speech to young infants and overspecified in speech to older infants while vowels exhibit an opposite trend. This hypothesis draws from the speech perception timeline that suggests that vowels are learned earlier than consonants. In support of this hypothesis, she found shorter VOTs in speech to 3-month-old infants (Sundberg and Lacerda 1999) but longer VOTs in speech to 11-14-month-old children (Sundberg 2001). However, Englund and Behne (2006) did not find support for this hypothesis based on infants' age as described above.

Another potential reason for the difference between IDS and ADS may be unrelated to any modifications to the segmental characteristics of the input, but rather a consequence of the prosodic characteristics of each register. For instance, given that IDS is characterized by a higher pitch, larger pitch contours and slower speech rate (see Section 2.1.3 for a description of prosodic features), Cristia argues that these properties could affect the acoustic features of speech sounds in multifaceted ways. For example, a slower speech rate could cause some categories to be exaggerated or enhanced. Thus, speakers may not have the intention to clarify phonetic distinctions and enhancement may merely be the result of speaking slowly. To account for the limitations of previous studies, Cristia (2010) examines the fricative sibilant contrast /s/ and /ʃ/ because this contrast is more resistant to the effect of prosodic features in IDS. She also uses a 'perceptually-based' definition to account for phonetic enhancement. She recruited 55 female caregivers with a similar educational background. 32 of these had infants between the age of 12 and 14 months and 23 of them had infants around 4-6 months old. These two age groups were used in order to detect the age-related variation. In a sound-treated room, the caregivers were recorded talking to their infants about some objects that were chosen to elicit the target consonants and some other filler items. After that, the experimenter engaged in an interaction with the caregiver discussing the same objects to elicit ADS. The results provided evidence for consonantal enhancement in speech directed to 12-14-month-old infants, even when prosodic characteristics (e.g. speech rate and pitch) were controlled for. This suggests that phonetic enhancement is modified based on the listeners' age or perceived linguistic needs.

3.1.2.2 Studies on vowels

Another line of research investigating the acoustic properties of IDS segments has come to the conclusion that vowel space is enhanced in IDS compared to ADS (Khul *et al.* 1997; Uther *et al.* 2007; Cristia and Seidl 2014; Hartman *et al.* 2017; See Chapter 4 for more details on how vowel space is related to speech perception). An enlarged vowel space in IDS has been interpreted as an indication of hyperarticulation, a process that indicates that vowels in IDS are articulated more clearly than those in ADS (Miyazawa *et al.* 2017). An influential study by Kuhl *et al.* (1997) provid cross-linguistic evidence for the enhancement of the vowel space area in IDS. The study examined language input directed to infants in the US, Russia, and Sweden to find out if phonetic units are modified in IDS. They recorded natural conversations of 30 mothers speaking to their 2-5-month-old infants and these 30 mothers speaking to an adult native speaker. Words containing the point vowels /a/, /i/ and /u/ were extracted from the speech samples for analysis. Acoustic analysis was undertaken to measure the vowels' formant frequencies (F1, F2, and F3) and pitch frequency at three points (onset, center and offset of the vowel). The study found that mothers across all languages enhanced the phonetic units in their speech when talking to their infants by producing more extreme vowels (Figure 3.1). The results indicate that mothers did not merely increase all vowel frequencies, but rather selectively increased or decreased formant frequencies to expand the acoustic space covering the vowel triangle. This vowel enhancement was not found in the mothers' speech to the native adult.

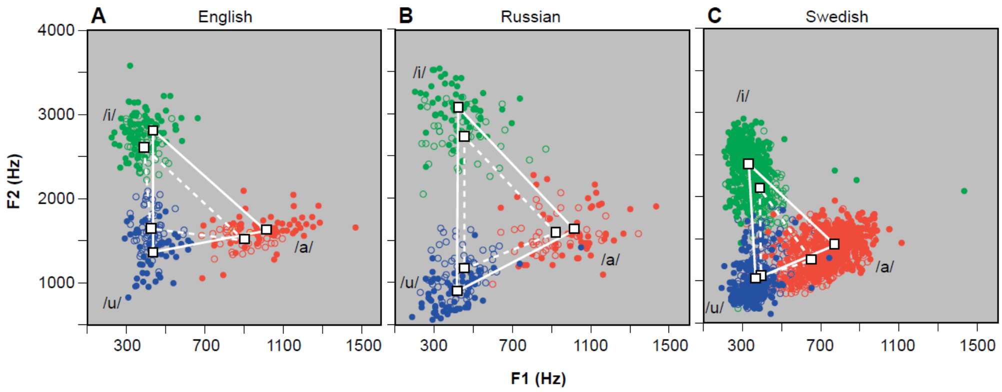


Figure 3.1 Cross-linguistic evidence of the expansion of the vowel space area in IDS compared to ADS in Khul et al.'s study (1997) (the dashed line represents ADS while the solid line represents IDS)

Whether an expanded vowel space is beneficial for infants or not, Kuhl *et al.* (1997) proposes some ways in which infants can benefit from an expanded vowel space in IDS. First, she claims that vowels become more distinct when the vowel space is acoustically stretched. Evidence for this comes from studies that show that the language capacity of language-delayed children improves when listening to speech in which between-category acoustic differences are expanded. Second, mothers hyperarticulate their vowels to produce instances that go beyond those produced in normal adult conversation. Research shows that hyperarticulated vowels present "better instances" of vowel categories when perceived by adults as well as children. Some studies attempt to find a link between mothers' vowel space expansion in IDS and their infants' discrimination ability. For instance, Liu *et al.* (2003) provides the first evidence of a positive correlation between the acoustic quality of maternal speech and infants' discrimination ability of phonetic segments. They tested vowel enhancement in the speech of 32 Mandarin-speaking mothers speaking to their 6-to-8- or 10-to-12-month-old infants, and compared these to the same mothers speaking to an adult. They found that mothers expanded the vowel space when addressing their infants but not when interacting with the native-speaking adult. They also used a Head-Turn experiment to test infants' speech discrimination when they were exposed to IDS vs. ADS. The results indicated that infants performed significantly better when exposed to IDS as opposed to ADS. Another study by Song *et al.* (2010) found that 19-month-old infants were able to look at a target object faster when they listened to IDS that exhibited exaggerated vowel space than when they listened to ADS. Both studies come to a similar conclusion which suggests that expanded vowel space in IDS facilitates infants' word recognition.

Despite evidence on expanded vowel space in IDS, not all studies examining the acoustic properties of IDS reported enhancement of vowel categories. For example, Englund and Behne (2006) did not find enhancement of point vowels in Norwegian mothers' speech to their infants. Likewise, a cross-linguistic study of the speech of British, French and Japanese mothers to their 6-to-22-month old children did not find the expansion of the vowel space in IDS, but rather reduction (Dodane and Al-Tamimi 2007). Instead, the vowel triangle in Dodane and Al-Tamimi's study was found be more expanded along the F1 dimension, suggesting more open vowels. The authors claim that producing more open vowels can make these categories more salient and easier to extract as the children were at a stage where they might benefit from matching articulatory gestures and sounds from speech. Furthermore, they suggested that mothers could be imitating their baby's vowels which are produced with a smaller vocal tract that triggers higher F1s. Similarly, Benders (2013) did not find an expansion of vowel space in the speech of Dutch mothers to their 11- and 15-month old infants. Vowel space was significantly shorter in IDS. Vowel formant frequencies (F1, F2 & F3) were found to be generally higher in IDS. Benders argued that raised formant frequencies in IDS are the result of positive affect that can be caused by mothers' smiling.

According to Wang *et al*. (2019), whether modifications in IDS occur or not may depend on aspects other than the nature of IDS. These aspects include infant's age, language typology, methodologies used to collect data or measure formants and culture.

3.1.3 Prosodic adaptations in speech to children

While research focusing on segmental aspects of IDS shows inconsistent results with regards to whether IDS is hyperarticulated or not, most research on IDS has agreed that prosodic features in mothers' speech to their infants are exaggerated. Cross-language research has shown that IDS has special prosodic features such as higher pitch, wider pitch excursions and contours, exaggerated intonation, higher intensity, longer pauses, shorter utterances and slower speech rate (Fernald and Simon 1984; Fernald *et al.* 1989). Infants under 9 months (Dunst *et al.* 2012) as well as children aged 12 and 16 months (Segal and Newman 2015) have been found to show a preference for IDS partly because of its prosodic properties. Researchers observe that mothers use exaggerated prosodic contours to direct their infants' attention and control their arousal level (Fernald and Khul 1987; Cooper *et al.* 1997). Segal and Newman (2015) claim that increased attention could be a source for language learning especially with the presence of other competing information sources that could encourage the infants' attention as well. Research on IDS has likewise shown that IDS is distinctive in the exaggerated pitch range and the intonation it exhibits compared to ADS. It is also believed that increased pitch in IDS serves an emotional role, especially that it has been compared to the non-emotional ADS (Fernald and Khul, 1987; Knoll *et al.* 2006).

In ADS, focus words are treated differently from a prosodic perspective. For instance, studies confirm that speakers use intonational prominence to stress information that is new in the context of previous discourse (e.g. Halliday 1967; Chafe 1974). Research that examine the acoustic correlates of linguistic focus show that increased duration and *f*0 are used as a means to highlight new words compared to background information (Brown 1983; Eady and Cooper 1986). Fernald and Mazzie (1991) argue that if prosodic prominence cues to sentence focus are useful to adults, they must be even more beneficial to infants. New words are seldom new in an absolute sense for adults whereas infants are constantly exposed to novel vocabulary and words they have not yet learned. Additionally, infants do not have the ability to exploit contextual details to process ambiguous sentences like adults. Hence, prosodic cues that increase the salience of words should be especially useful to preverbal infants.

With regard to speech rate of IDS, studies have reported shorter utterances, longer vowels, longer pauses, and final syllable lengthening (e.g. Fernald *et al.*1989; Swanson *et al.* 1992; Uther *et al.* 2007; Cristia and Seidl 2014). A cross-language study by Fernald *et al.* (1989) found that mothers' and fathers' speech to their preverbal infants is characterized by shorter utterances and longer pauses compared to their speech to adults. In another study, Swanson *et al.* (1992) found that vowels of content words, especially those in final phrase position, in the speech of 15 English-speaking mothers to their 1;6-to-2;4-month-old children were much longer than vowels in ADS. Similarily, Church et al. (2005) reports that mothers exaggerated overall utterance-final syllables in IDS compared to ADS, which significantly affected speech rate. Another study, by Ko (2012), explores the role of IDS in language acquisition by examining how mothers' speaking rate change over the course of their children's language development. For analysis, the study used data from the CHILDES database of 25 British and American mother-child pairs. The speaking rate was calculated using the number of words per second in utterances that did not exceed 10s, as these might contain long pauses. The results showed that IDS speaking rate changed in a nonlinear fashion in the pre-verbal stage of the children but shifted around the stage of the children's onset of speech production. There was then acceleration in the speech rate until the age of 2 before it started varying again in later stages of development. The author concluded that the results appear to provide support for the view that IDS is modified based on the child's linguistic needs at a particular point in language development. Table 3.1 summerizes the phonological, acoustic and prosodic features of IDS.

Phonological	Acoustic	Prosodic
Consonant substitutions	Consonant enhancement (some studies did not find this enhancement)	Slower speech rate (longer vowels and syllables)
Consonant cluster reduction	Vowel space expansion (some studies didn't find this expansion)	Exaggerated *f*0, *f*0 contours and wider pitch
Phonological reduction (in some studies)		Higher intensity

Table 3.1 Phonological, acoustic and prosodic features of IDS

3.2 Clear Speech

Talkers tend to use a 'clear speech' style, consciously or unconsciously, when speaking in difficult communication conditions such as speaking in a noisy or reverberant environment or talking to a hearing-impaired listener or a non-native listener. Clear speech is usually described by comparing it to conversational or plain speech which is produced under typical conditions when talkers make no special effort to speak clearly (Uchanski 2004). Studies on speech clarity have been mostly concerned with investigating the effect of clear speech on intelligibility and perception. Previous studies that focused on communication between normal hearing speakers in noise conditions found that clear speech was used to achieve improved intelligibility (e.g. Snidecor *et al.* 1944; Tolhurst 1958, 1957). Other studies found a similar intelligibility advantage of clear speech for the hearing-impaired listeners in both quiet and noise environments (e.g. Pichney *et al.* 1985; Uchanski *et al.* 1996). The main focus of the previous studies of clear speech has been on communicative conditions in which speech communication is challenged by some sort of noise or interlocutor's hearing loss that impede the listener's access to the speech signal. Bradlow and Bent (2002) examine whether clear speech will have a similar intelligibility advantage for non-native listeners. According to the researchers, NNSs have speech perception problems related to their limited access to the sound system and higher linguistic structures of the target language. 32 non-native listeners and a control group of 32 native listeners were presented with naturally produced English sentences (conversation vs. clear). The study found that clear speech had a very small effect on improving the intelligibility of non-native listeners, whereas it had a greater benefit for improving the intelligibility of native listeners. The authors explain this in terms of the fact that clear speech is "native-speaker oriented" and therefore native speakers get the most benefit from it as they are experienced with the phonological and acoustic structure of the target language compared to non-native speakers (p.281).

Let us now focus on the acoustic, prosodic and phonological properties of the speech signal, ignoring other characteristics that are likely to account for speech clarity such as simpler syntax and vocabulary as they are not relevant to the present study. The difference in the physical characteristics of clear speech and conversational speech has been examined based on three levels:

- Global: changes in intensity, speaking rate, pauses, fundamental frequency, long-term spectra;
- Phonological: feature changes, deletions or insertions of phonemes;

- Phonetic: spectra and durations, formant frequencies of vowels, VOTs of consonants, amplitudes.

Speaking rate is a prominent global talker feature and has been reported to distinguish clear speech from conversational speech (Picheny *et al.* 1986; Krause and Braida 1995). Also, it has been found to distinguish IDS from ADS as discussed in section (3.1) and FDS from ADS (Section 3.3). Studies that investigate this global feature have provided a number of factors that contribute to variation in speaking rate such as frequency and length of pauses, number of words or syllables per minute, length of final syllables as well as speech segment duration. In a series of experiments to examine whether intelligibility of hearing-impaired listeners is improved by speaking clearly, Pichney *et al.* (1985, 1986, 1989) establish that speaking rate is one robust actor that distinguishes clear speech from conversational speech. One study in this series (Picheny *et al.* 1986) was concerned particularly with comparing acoustic properties of clear speech against conversational speech. Three male speakers were asked to produce 50 non-sense sentences once in a clear speech mode and another in a conversational speech mode. The average speaking rate was determined by three factors: number of words per minute (w/m), pause duration and phoneme duration. The study found that the range of speakers' speaking rate in conversational speech was much higher (160-200 w/m) compared to clear speech (90-100 w/m). Clear speech was found to have longer pauses than conversational speech. Also, the duration of speech segments was increased in clear speech which contributed to a slower speaking rate according to the authors. Pitch or fundamental frequency ($f0$) is another global feature that has been found to change when speakers are asked to speak clearly vs. naturally. Clear speech is generally described to have a higher pitch and a wider range in $f0$ than conversational speech (Uchanski 2004).

Clear speech research has also documented acoustic modifications at various levels of the speech segment. Pitcheny *et al*. (1986) report an increase in VOTs for voiceless plosives in word-initial positions in the clear speech of two of his participants. In addition, the duration of some speech segments such as fricatives, nasals, and semivowels is found to be increased in clear speech. The authors also measured the segmental power (the energy of the segment at a particular point in time) of different classes of speech segments. The results showed a higher intensity for consonant sounds especially plosives in clear speech. At a phonological level, the study found that vowels in clear speech were unmodified while they were reduced in conversational speech. Clear speech was also characterized by stop burst releases in word final position. In addition to these findings, Bradlow *et al.* (2003) note that clear speech also has less alveolar flapping than conversational speech in a study that investigated the effect of clear speech on American children with learning disabilities. Clearly produced English fricatives are further investigated by Maniwa *et al.* (2009) who found that fricatives in clear speech were significantly longer in duration, have a higher spectral mean, lower skewness and higher peak frequency.

Vowel modifications have been recurrently investigated in clear speech research. The most substantial vowel changes reported in this research are an increase in vowel duration (Pitcheny *et al*. 1986; Ferguson and Kewly-Port 2002) and vowel space size (Pitcheny *et al.* 1986; Bradlow *et al.* 1996; Ferguson and Kewly-Port 2002) for clear speech relative to conversational speech. Ferguson and Kewley-Port (2002) examine among several other experiments, the acoustic characteristics (spectral, dynamic and durational) of 10 vowels in /b*V*d/ context in clear vs. conversational speech. The test words that were embedded in 12 sentences were elicited from a male speaker who read the sentences in two styles: (1) clear as though he were talking to a hearing-impaired person and conversational using his normal style. Based on average steady-state results of F1 and F2, vowel space area of clear speech was significantly larger than that of conversational speech. F1 was found to be significantly higher in clear speech compared to conversational speech regardless of the vowel category. Higher F1 indicates an increase in vocal effort caused by speaking clearly (Lienard and Dibenedetto 1999). Changes in F2, however, were found to depend on the vowel (Ferguson and Kewley-Port 2002). F2 of all front vowels was significantly higher in clear speech than in conversational speech, whereas, F2 of back vowels was overall lower in clear speech, but this result was only significant for the /ʌ/ and /ʊ/. Vowels were also found to be more dynamic in clear speech than conversational speech based on vector length results (F1xF2 space). With regard to duration, vowels in clear speech were significantly longer (2.1 times longer) than vowels in conversational speech. Table 3.2 illustrates the most investigated features of clear speech.

Acoustic	Prosodic
Consonant contrast enhancement (clearer consonants)	Slower speech rate (longer vowels and consonants)
Vowel space expansion	Exaggerated *f*0, Higher intensity

Table 3.2 Acoustic and prosodic features of clear speech

3.3 Foreigner-Directed Speech

The study of how people modify their speech when addressing foreigners started as early as the nineteenth century when this kind of language was discussed in relation to pidgin lingua francas (Ferguson 1981). SLA research that focused on input as an independent variable discussed three simplified registers that are used in NS-NNS interactions: foreigner talk, teacher talk and interlanguage talk (Beebe and Giles 1984). Ferguson (1971) coined the term 'foreigner talk' and described it as a register used by native speakers of a language when speaking to older hearers who are assumed to have less knowledge of the language than the native speaker. 'Foreigner-directed speech' is a more recent term that has been used mainly in studies that invstigate the acoustic and prosodic characteristics of speech addressed to L2 learners (see Bobb *et al.* 2019 for a review, Section 2.3.2). Foreigner talk and the other codes just mentioned were considered as comparable in some ways to other speech codes, such as baby talk, speech directed to hearing-impaired listeners and pidgins (Freed 1981; Wesche 1994; Ellis 1994). The basis for this assumption is that they are all characterized by the 'linguistic insufficiency' of the listeners (Freed 1981:20). However, according to Freed, both registers (baby talk and foreigner talk) also differ in some attributes; for example, children are cognitively restricted, immature and are of reduced status in relation to caregivers, whereas, foreigners are cognitively equal to adult native speakers and are not of inferior status in relation to native-speaking adults (ibid). Researchers interested in examining NS modifications in instructional settings used the term 'teacher talk' (Zuengler 1991; Routman 2002). Some SLA studies in instructional settings have been interested to investigate whether simplified input facilitates communication in a given interaction and whether it has a longstanding effect on the acquisition of the L2 (Beebe and Giles 1984; Ghorbanian and Jabbarpoor 2017). Teacher talk will not be examined here as the present study mainly focuses on foreigner talk.

According to Dela Rosa and Arguelles (2016: 46), comparing foreigner talk to other simplified registers indicates that it “reflects universal processes of simplification, knowledge of which constitutes part of a speaker's linguistic competence”. In the same vein, Hatch (1983) suggests that foreigner talk is characterized by three aspects: (1) regression, (2) matching and (3) negotiation. *Regression* indicates that NSs move back through their own stages of linguistic development until they find a suitable level. *Matching* refers to when NSs evaluate a NNSs’ existing interlanguage status and accordingly imitate forms they detect in it. Finally, *negotiation* reflects how NSs simplify and clarify their speech according to the feedback they obtain from the NNSs.

A considerable amount of research has documented the linguistic and conversational adjustments made by native speakers (NS) of a language when addressing non-native speakers (NNS) (e.g. Ferguson 1975, 1971; Meisel 1977; Arthur *et al.* 1980; Freed, 1981; Long 1983a; Long 1983b; Gass and Mackey 2007; Gass 2009; Rodriguez-Cuadrado 2018; Zuraida and Fitri 2019). Early work on foreigner talk in the 1970s and 1980s focused on one dominant feature of this register which was ungrammaticality (Long 1983). Ferguson (1971) states that the nature of ungrammaticality in foreigner talk is manifested in different processes such as omission, expansion, and replacement. Long (1981) proposes certain conditions under which ungrammatical foreigner talk manifests itself, including (1) the NNSs have low proficiency in the L2, (2) the NNSs are perceived of having lower status by the NSs (3) NSs have prior experience with the NNSs and (4) NSs and NNSs are involved in a spontaneous interaction (Long 1981). Examples of ungrammatical speech in foreigner talk come from different studies. For instance, Snow *et al.* (1982) report that foreigner talk involved imitation of foreigners’ errors in the Netherlands such as incorrect deletion of determiners and incorrect word order. Snow and Hoefnagel-Höhle (1982) claim that such distortions of the grammar do not occur in speech addressed to children, which can explain why adult second language acquisition is imperfect and less successful than that of first language acquisition.

Other studies from the same period report that speech addressed to NNSs is, in fact, a modified but well-formed version of the target language. For instance, some of the foreigner talk features observed in these studies are the use of syntactically fewer complex structures (Freed 1981), the use of shorter expressions (Arthur *et al.* 1980; Freed 1981) and the use of restricted vocabulary (Arthur *et al.* 1980). Long (1983) suggests that it is essential for second language researchers to identify which of the input elements are grammatical and which are not. This is because learners who end up fossilizing with a ‘pidginized’ version, as Long referred to it, of the target language are most likely those who often receive “restricted input, ungrammatical input or both” (Long 1983: 1979).

Based on a review of studies on foreigner talk, Zuengler (1991), notes that not all NSs use foreigner talk when interacting with NNSs, nor does foreigner talk, when used, contains all of the features just identified. According to Zuengler, one reason for the disagreement over the grammaticality of foreigner talk and for when and why it is employed is the absence of an overriding theory of foreigner talk (ibid). The variability in the foreigner talk features could be attributed to different factors. Snow *et al.* (1981) speculates that among the factors that may lead to variability in foreigner talk findings is the foreign addressee, the perceived language competence of the foreign addressee, the native speaker and the nature of the conversation. Foreigner addressees who are perceived as intelligent, high status and socially equal to native speakers are addressed with the least extreme characteristics of foreigner talk unlike low-status foreigners (Snow *et al.* 1981; Zuengler 1991). Moreover, foreigners who are less foreign accented and have more second language production skills may trigger less foreigner talk features. For instance, in a study that examined the interactions between Dutch-speaking municipal employees and non-Dutch speaking foreigners in Netherlands, Snow *et al.* (1982) found that foreigners who produced longer and more complicated utterances received relatively long sentences back, whereas foreigners who produced simple and short sentences triggered short and simple utterances back.

Different NSs may also produce different characteristics and different amounts of foreigner talk. For instance, Snow *et al.* (1981) claims that NSs who have experience and long contact with foreigners are well aware of the type of adaptations that can ease communication with their foreign interlocutors, whereas NSs who do not have considerable experience with their foreign interlocutors may not know how to produce foreigner talk. Conversely, Gass and Varonis (1984) found in a study that NSs’ familiarity with non-native discourse is important in determining the changes that are required by the NS in interaction with their foreign interlocutors. A number of studies also reported that the most extreme characteristics of foreigner talk were used by NSs in conversations with their close foreign friends (Hatch *et al.* 1975; Katz 1977) or by second language teachers who had professional acquaintance with foreigners (Hatch *et al.* 1975; Chaudron 1978). Another source for the variability in foreigner talks features is the nature of conversation carried out by NSs and NNSs. Snow *et al.* (1981) report that in a study of Dutch foreigner talk they conducted in 1978, longer conversations initiated by the foreigners' triggered more foreigner talk features than shorter ones.

Now that we know the nature of foreigner speech characteristics discussed in early research, the following sections will be devoted to work on the phonological, acoustic and prosodic features of FDS, the main focus of the present study.

3.3.1 Phonological adaptations in FDS

A few early studies described foreigner talk from a phonological perspective. Old studies relied mostly on auditory observations to account for the phonological adaptations in FDS. According to Hatch (1979), foreigner talk is characterized with the following phonological features: clear articulation, slow rate, release of final stops, more heavily-voiced final stops, insertion of glottal stops before words starting with vowels, fewer reduced vowels, longer pauses, extra volume and exaggerated intonation. According to Tarone (1980), these phonological characteristics are production strategies that the speaker using foreigner talk employs in order to ease the processing of language for the learner. Several other studies report that teacher talk is characterized with clearer articulation and slower speech rate (e.g. Henzle 1979; Chaudren, 1988). Henzl (1975) also found that speech rate in foreigner talk is affected by the proficiency level of the NNS. In his study, Henzl (1975) report that NSs slow their speech rate to a greater degree when interacting with beginner NNSs than wheh addressing advanced NNs

Other types of phonological modifications in foreigner talk come from a study discussing the use of English foreigner talk by second generation children when addressing their German-, Dutch-, Italian-, Hungarian-, Yiddish-, Maltese- or Greek-speaking parents and grandparents who immigrated to Australia. Findings indicate that foreigner talk is phonologically and syntactically marked (Clyne 1981). Some of the phonological properties of this code reflect features of the languages of the addressees. General phonological features of foreigner talk that were reported in this study include the following: Monophthongization (e.g. /əu/→[u:], /ʌu/→[o:]), shortening (e.g. /i/→[ɪ]), Backing (/æə/→[au]), and devoicing (e.g. /z/→[s], /d/→[t], /ð/→[d] or [s], /θ/→[t] or [s], /v/→[f]) (pp. 73-75). As to why children of immigrants modify their speech phonologically when speaking to their parents, the informants reported that this was for ease of communication. Clyne, however, suggests that this reason was surprising given that the addressees seemed to have no difficulty involving in conversations with native Australians. Alternatively, the author suggests that another explanation for the use of this code could be that the speakers were identifying with their addressees or continuing to use a variety of English they originally used themselves at some point in their lives.

Katz (1981) found similar phonological features to the ones reported by Clyne in a study that investigated child-to-child foreigner talk. The study examined the speech of an American five-year-old-girl (Lisa) talking with her age-peered Hebrew native-speaker playmate (Tamar) longitudinally in a natural play setting. Tamar had no previous exposure to English and had been acquiring the target language without formal instruction. The most striking phonological aspects the study found for Lisa's foreigner talk were specific, low- level features that were used more frequently than global features (e.g. exaggerated intonation, slower and careful speech and syllable-timed utterances). For example, Lisa's foreigner talk was characterized by frequent vowel de-tensing and de-laxing (e.g. /e/→[ɛ], /i/→[ɪ], respectively), excessively nasalized vowels, de-sulcalization (e.g. /əɹ/→[ʌ], [ə], [ö], [ʊ]), /r/→tap, trill, [w], [l], [n], initial non-aspirated stops and monophthongization of diphthongs. Katz report that, in contrary to his predictions, the phonological features of Lisa's foreigner talk were increasing over time compared to the morpho-syntactic features. The basis for Katz's initial predictions was that as a 'foreign accent' of a foreigner's L2 might be the most prevalent character of their speech, native speakers are expected to grow more familiar with this accent, but at the same time start to notice the structural deviations of a foreigner's speech from the target language. For this reason, the phonological peculiarities of foreigner talk become less prominent over time or remain constant, whereas the morpho-syntactic patterns become increasingly manifest. An alternative hypothesis that Katz provided to explain the rather unexpected behaviour of Lisa's foreigner talk is that since an L2 learner's morpho-syntax is deemed to improve over time in contrary to L2 phonology which is more resistent to improvement, foreigner talk ought to reflect these facts if it is sensitive to the structure of the L2 learner's output.

In another account, Nelson (1992) reports on the phonological features of foreigner talk as used by her father (a physician) while talking over the phone to his native and non-native English-speaking patients. Among the adaptations used were greater speed, more careful pronunciation and more syllable-timed stress, less consonant cluster reduction. For instance, Nelson describes vowels as produced “in their pure, stressed form” (p.26).

3.3.2 Acoustic and prosodic adaptations in FDS

While earlier work on foreigner talk focused more on grammaticality and auditory observations of phonological features, subsequent research has sought to investigate the acoustic and prosodic properties of foreigner talk (referred to as FDS in this research) more thoroughly. This research has primarily drawn parallels between FDS and IDS or speech directed to hearing-impaired listeners, showing that foreigners might get exposed to similar acoustic and prosodic modifications due to their limited competence of the lower and higher linguistic structures of the target language and apparent difficulty accessing the linguistic code of the native speakers (Bradlow and Bent 2002). Despite its importance, this area of research remains understudied and relatively little is understood about the acoustic and prosodic features of FDS. This section reviews the few studies on FDS available in the literature presented in chronological order.

In a first attempt to investigate the prosodic features of FDS, Biersack *et al.* (2005) examine the prosodic properties of IDS and FDS in order to find out whether FDS is a ‘derivative’ of IDS. They asked 12 female native English speakers to use a map and describe the route to imaginary interlocutors: an adult friend, a two-year-old child and a foreigner who is new to the country and has just started learning the language. Each participant used three map versions with different landmarks to avoid repetitiveness in the descriptions and they started the description using one of two calibration sentences. The study measured three prosodic features: F_0 maxima, pitch range and speech rate (indexed as the number of syllables per second, vowel durations and pause durations) mainly of the calibration sentences.

Results show that compared to ADS, speakers increased their pitch range when addressing a child but not a foreigner. FDS was narrower in pitch than ADS, though this trend was insignificant. The authors concluded that the ‘attention-eliciting’ prosodic properties speaker use when addressing a child are absent in FDS. Results also revealed that participants slowed their speech rate (measured by number of syllables per second) significantly in FDS than in IDS or ADS. Moreover, FDS had longer pauses than ADS or IDS. Vowel lengthening was significantly evident in IDS. Similarly, FDS exhibited longer vowels compared to ADS but this was insignificant. The researchers concluded that speakers tend to adjust the prosody of their speech depending on the needs of the interlocutor. The authors concluded that this variation in the use of different modification strategies of speech rate in IDS and FDS is adjusted to suit the needs of the different interlocutors. To this end, the authors claimed that they provided evidence contrary to the belief that FDS is a “mere derivative” of IDS.

Although the previous study had the advantage of measuring pitch range and speech rate using a calibration sentence which was identical across speech registers and thus obtains results that were comparable across speakers, using participants imagining talking to hypothetical listeners has the disadvantage of eliciting interactions that are unspontaneous and contrived. Moreover, the findings of this study might not be comparable to other studies that used natural interactions.

Uther *et al.* (2007) took the first initiative to investigate whether vowel hyperarticulation occurs in FDS. One main motivation for their study was that since IDS serves a didactic role, then speech to other addressees who are precieved to require ‘linguistic instruction’, such as foreigners, should also serve a didactic role (p.3). Uther *et al.* also sought to uncover whether vowel hyperarticulation in FDS is independent of the attentional and affective characteristics of IDS. They recorded the speech of ten Southern British-English mothers in naturalistic interactions with their own infants aged between 4 months and one year, a British adult confederate, and one of two Chinese adult confederates with a perceived foreign accent. Three dependent variables were included in the study to find answers to the previous questions: pitch, vowel space of each of three corner vowels, /i/, /u/, /a/ and voice affect. In a quiet room in their own house, mothers were given three toys (a sheep, a shoe and a shark) and were asked to communicate with their interlocutors using these toys in three different consecutive sessions. The target vowels were analyzed in Praat for pitch (indexed as mean F0) and vowel space (F1/ F2). Voice affect was evaluated by independent raters.

The study found that both IDS and FDS exhibited significant vowel hyperarticulation compared to ADS, and that vowel space of FDS and IDS was similar. The authors concluded that this provides strong evidence that vowel hyperarticulation and its role in facilitating language learning in IDS is generalizable to FDS. In other words, IDS and FDS are both didactically oriented. Furthermore, findings revealed that the degree of hyperarticulation with regard to the front and back vowels in FDS was not as robust as that reported in previous research. This was attributed to the mothers' Portsmouth dialect that differs from the English dialects of mothers who participated in previous studies. To ensure that vowel length did not trigger vowel hyper-articulation, Uther *et al.* measured vowel duration. They found that vowels in IDS were significantly longer in duration compared to those of ADS. However, vowels of FDS were insignificantly shorter in duration than those of ADS. Hence, the authors concluded that as patterns of vowel duration did not reflect those of vowel space, vowel length played no role in vowel hyperarticulation. The study also found that FDS showed lower pitch and less positive affect than IDS. However, FDS was equal in pitch and significantly lower in positive affect than ADS. From this, the authors concluded that heightened pitch and positive affect are not linguistic features of speech modification as IDS and FDS both exhibited vowel hyper-articulation but varied in pitch and affect rating.

Scarborough *et al*. (2007) investigate the acoustic characteristics of FDS and compare them to those of ADS as well as those found in other speech registers. They also examine whether speech directed to real interlocutors is comparable to that directed to imaginary ones. They recruited 10 American English speakers (7 males and 3 females), a female American English confederate and a Mandarin Chinese confederate with less than 3 years of residence in the U.S. Each participant was asked to use maps and describe their route to 4 interlocutors in 4 experimental conditions: an imaginary foreigner, an imaginary native speaker, a real native speaker confederate and a real foreigner confederate. The order of the conditions was mixed across participants. Data analysis was carried out using Praat. Measurements included F1 and F2 of the point vowels /i/, /u/ and /a/ at the midpoint, vowel space, vowel duration and speech rate (indexed as word per second and segment per second as well as average word and segment durations). Only vowels of stressed syllables were reported in this study.

Results revealed that speech rate was significantly slower in FDS compared to ADS, and that speech to imaginary interlocutors was even considerably slower than speech to real interlocutors. The results also showed that in FDS, speakers produced significantly longer vowels than in ADS. Vowel duration was also considerably longer in speech to imagined interlocutors than that to real interlocutors. Vowel space (statistically represented by the distance between vowels) was found to be more expanded in FDS compared to ADS and further expanded in speech to imagined interlocutors than in speech to real interlocutors. The authors conclud that speech directed to foreigners is comparable to speech directed to hearing impaired listeners or speech produced in noisy settings. They also establish that although patterns of adjustments were similar in the FDS conditions, speech to real interlocutors was not acoustically indistinguishable from speech to imaginary interlocutors. Therefore, they suggest that future studies need to consider authentic interactions when designing their methodologies.

The previous study provides some insights as to how using hypothetical listeners to elicit FDS could provide findings that are misleading. However, the study overlooked the ways in which males' and females' speech could vary acoustically (Barney 1952; Fant 1976). Using speech samples from males and females could generate data that is incomparable across speakers.

In another study that aimed to evaluate the validity of using hypothetical interactions in speech research, Knoll and Scharrer (2007) replicate Uther *et al.*'s (2007) study but using imaginary interlocutors instead of real ones and laboratory setting instead of the home environment. They recruited 10 British female students who had a Southern English accent. Speakers were instructed to imagine talking to an infant, a British adult and a foreigner in different sessions carried out in a laboratory. The same data collection tools and linguistic variables as in Uther *et al.* were used. Unlike what Uther *et al.* found, the results from this study indicate a significantly narrower vowel space in IDS compared to both adult conditions. The authors explain this in terms of a lack of experience from the part of the speakers and the absence of feedback from the imaginary child. FDS and ADS were statistically comparable in terms of vowel space size. Unlike Uther *et al.*, this study did not find any significant difference in vowel duration between the three registers. With regards to pitch ($f0$), IDS was found to be significantly higher in pitch than ADS but mean pitch of IDS was lower than that reported in Uther *et al.*, IDS and FDS were not comparable in terms of pitch and the difference in pitch between FDS and ADS approached significance ($p=0.054$). Similar to Uther *et al.*'s findings, IDS was significantly higher in vocal positive affect compared to the adult conditions and FDS was lower than ADS in positive affect but this was not significant.

Based on these results, Knoll and Scharrer (2007:1417) conclud that their findings mostly do not approximate those of Uther *et al.* (2007) except for vocal positive affect. They suggest that this was the case probably due to the absence of "two-way dynamic feedback between speaker and listener" which is crucial for generating suitable speech adjustments; thus, they recommend that speech to real interlocutors should not be underestimated especially in research that requires its results to be generalizable to the real world.

As previous FDS examined NSs' speech to foreign accented speakers regardless of their characteristics, Kangatharan (2014) examines the role of interlocutors' appearance and accent on triggering vowel hyperarticulation in FDS. Fivty-two white British adults engaged in a spot the difference task with one confiderate from four groups: (1) Native looking and native sounding (NLNS), (2) Native looking and foreign sounding (NLFS), (3) Foreign looking and native sounding (FLNS), and (4) Foreign looking and foreign sounding (FLFS). She examines adaptations in vowel space, mean *f*0, intensity, vowel and word duration of FDS. In line with Uther *et al. (2007),* the study found that speech addressed to foreign accented speakers was hyperarticulated through expansion of vowel space. Also, NSs produced longer vowels and words in FDS compared to ADS. The results revealed that vowel hyperarticulation was attested in the NLFS and FLFS conditions, in which the interlocutor was foreign sounding. No effect of appearance was attested for this property. Similarly, word and vowel length increased significantly in the foreign sounding conditions but not in the native conditions. The author justifies the previous two findings in terms of the H&H theory which proposes that speech will be adapted based on the listeners' needs in communication. On the contrary, mean intensity was found to be significantly higher in speech addressed to foreign looking interlocutors irrespective of accent. The author explains this in terms of previous findings that indicate that intensity is not an essential property in speech addressed to non-native speakers (e.g Sommers *et al.* 1994). Additionally, in line with previous research, mean *f*0 was the same in all conditions.

In a more recent study, Knoll *et al.* (2015) examine vowel space expansion, mean *f*0, mean duration, *f*0 range and intensity in FDS compared to ADS and hearing-impaired directed speech (HIDS). To rule out how listeners perceive these speech types, they also examined clarity and positive vocal affect. Using the DiapixUK spot the difference task, ten Scottish females were recorded interacting with an adult Scottish female, one of three foreign confederates with a noticeable accent and 2.5-year residence in Scotland as well as a hearing-impaired female. Acoustic analyses carried out in Praat focused on three corner vowels /i/, /u/, /a/ extracted from speech samples. The results revealed that there was no significant difference between the three speech types with regard to mean intensity, mean duration or mean *f*0 range. FDS and HIDS had significantly higher mean *f*0 than ADS. Vowel space area was significantly maximized in HIDS compared to ADS but not in FDS. Also, twenty British participants rated HIDS to have a significantly more positive vocal affect and more intelligibility than ADS or FDS. FDS was not rated to be different from ADS with regard to these attributes.

Table 3.4 summarizes the previous studies and presents results concerning FDS as it was compared to ADS. Inconsistent results from the previous studies with regard to the expansion of vowel space area, vowel duration and *f*0 called for further investigation of FDS. Hence, the current study was motivated to further explore features examined in FDS including vowel space area, vowel length, intensity and *f*0 as well as expand this research by examining whether consonant contrasts are enhanced in such speech settings. Additionally, all aforementioned studies on FDS examine speech addressed to unfamiliar foreign accented adults. We know little of whether the same properties will occur when NSs address familiar foreign accented adults, such as when employers address their foreign domestic helpers. One more gap in the previous FDS research is the lack of examination of the role of the foreign interlocutors' linguistc developmental level in triggering FDS. Hence, the present study will explore this by relying on LoR of the FDH participants as an index to the learners' L2 experience.

The three registers so far discussed (IDS, clear speech and FDS) share some common features such as expanded vowel space and slower speech rate. However, FDS differs in other acoustic and prosodic aspects. Table 3.3 presents a comparison of the most prominent features of these registers.

Infant directed Speech	Clear Speech	Foreigner Directed Speech
Expanded vowel space (sometimes no expansion)	Expanded vowel space	No difference or expanded vowel space
Slower speech rate	Slower speech rate	Slower speech rate
Longer vowels	Longer vowels	No difference or longer vowels
Higher *f*0	Higher *f*0	No difference or higher *f*0
Higher intensity	Higher intensity	No difference in intensity
Enhancement of consonant contrasts (sometimes no enhancement)	Enhancement of consonant contrasts	Not Investigated

Table 3.3 Comparison of the acoustic and prosodic features of IDS, clear speech and FDS

Study	Interlocutors	Foreigners' proficiency level	Variables examined	Task	Results (FDS compared to ADS)
Biersack *et al.* (2005)	imaginary	---	$f0$ maxima, pitch range, speech rate (syllables/s, vowel duration, pause duration)	Map	-No significant difference in pitch range (FDS narrower in pitch). -FDS had significantly longer syllables and pauses. -No significant difference in vowel duration (FDS had longer vowel duration).
Uther *et al.* (2007)	Two Chinsese confederates	Detectable foreign accent	Mean $f0$, vowel space, voice affect, vowel duration	Natural conversation using toys at home.	-Significant pitch lowering in FDS. -Significant vowel expansion in FDS. -No difference in vowel duration (FDS had a shorter vowel duration)
Scarborough *et al.* (2007)	Imaginary and a Mandarin Chinese confederate	Less than three years of residence in the U.S.A	Vowel space, vowel duration, speech rate (word/second, segment/second, average word and segment durations)	Map	-Significant vowel space expansion in FDS (imagined more than real). -Significant word and vowel duration (longer) in FDS (slower speech rate).
Knoll and Scharrer (2007)	Imaginary	---	Mean $f0$, vowel space, voice affect, vowel duration	Imagined conversation using toys in a laboratory	-No significant difference in vowel space or vowel duration. -No significant difference in pitch (FDS was lower in pitch)
Kangatharan (2014)	One white British confederate, one white European, two Asian speakers.	Foreign accented	Mean $f0$, vowel space, voice affect, vowel & word duration, mean intensity	DiapixUK spot the difference task	-Significant vowel space expansion in FDS. -Significant vowel and word lengthening in FDS. - No significant difference in pitch. -Significant increase in intensity in the native condition.
Knoll *et al.* (2015)	Three Estonian and Czech foreigners	2.5 years residence in Scotland and noticeable accents	Mean $f0$, $f0$ range, mean duration, mean intensity, vowel space, vocal affect and clarity	DiapixUK spot the difference task	-No significant difference in mean intensity, $f0$ range, mean duration or vowel space. -A significant difference in mean $f0$(FDS was higher in pitch)

Table 3.4 Summary of FDS studies available in the literature

3.4 NS-NNS Communication and Interaction: Theoretical Frameworks

Now that it is clear what the acoustic characteristics of FDS are, this section presents the most prominent models and theories that can be used to explain the nature and dynamics of NS-NNS speech and interaction patterns. Discussions of the linguistic adaptations in NS-NNS interaction patterns (whether it is foreigner talk or foreigner-directed speech) have been directly or indirectly supported by theoretical approaches from speech communication (hypo and hyper theory) and sociolinguistics (communication accommodation theory). Hence, the current study found it essential to include these models and use them to later inform the results of this study.

3.4.1 Communication Accommodation Theory

Communication Accommodation Theory (CAT) is a comprehensive framework that aims to explain the adaptations individuals make to "create, maintain or decrease the social distance in interaction" (Giles and Ogay 2007: 293). Adjusting one's communicative behavior in such a manner is referred to as "accommodation" in CAT terms. Central to the theory is the notion that individuals who engage in an interaction use some communicative strategies (e.g. convergence and divergence) to mark their attitudes towards one another and their social groups (Giles and Ogay 2007). *Convergence*, the most widely studied strategy, refers to how individuals change their communicative behavior in terms of a wide variety of linguistic (e.g. accent, speech rate), paralinguistic (e.g. pauses, vocal intensity) and nonverbal patterns (e.g. head nodding, smiling) such that they become similar to their interlocutor's behavior (Giles 1973; Giles *et al.* 2005; Giles and Ogay 2007). Inversely, *divergence* refers to the way in which speakers "accentuate" their speech and nonverbal dissimilarities between themselves and the interlocutors (Giles *et al.* 2005; Giles and Ogay 2007). A strategy similar to *divergence* is *maintenance,* whereby an individual perseveres his communicative behavior or style irrespective of the interlocutor's communicative behavior (Bourhis 1979). One of the principle distinctions in the *convergence/divergence* paradigm is whether these two strategies are "upward" or downward" in terms of their societal value (Giles and Ogay 2007). *Upward convergence* depicts an individual's use of the prestige features of an upper class interactant, whereas *downward convergence* refers to adjustments made towards a stigmatized or less prestigious speech variety (Dragojevic et al. 2016). Conversely, adopting a standard variety with a non-standard accented interactant is an instance of *upward divergence,* whereas emphasizing one's own nonstandard speech with a standard-accented interact is an instance of a *downward divergence* (ibid).

CAT also proposes two different functions for accommodation: cognitive function and affective function. The cognitive function refers to how discourse in a communication is managed and the extent to which speakers are understood (Gallois *et al.* 2005). The affective function concerns effects such as “managing social distance”, “maintaining social control as it relates to power or status differentials”, “identifying to others” and “maintaining face” (Giles and Gasiorek 2013: 5). Several adjustment strategies can be used to achieve these functions. First, *approximation* strategies are used when speakers adjust their verbal or nonverbal performance to be similar or different from their interlocutors. These are similar to the convergence and divergence communicative strategies mentioned earlier. Second, *interpretability strategies* are used when speakers adjust their speech and the way they deliver information when they wish to enhance comprehensibility for their audience. Likely forms of interpretability strategies include adjusting speech rate, tempo, prosody, volume, syntax and repetition (Giles and Coupland 1991). Third, *discourse management strategies* are used when speakers adapt their communication on macro level through, for instance, “topic selection”, “topic sharing” and “turn management” (Giles and Gasiorek 2013: 6). Finally, *interpersonal control strategies* are used when speakers adapt their communication to tackle the social dynamics of the interaction by, for example, acting in relation to power or status relations.

Zuengler (1991) proposes that a number of the dimensions of CAT, especially those used to explain interactions with the elderly can be used to explain the dynamics of NS-NNS interaction. This proposal is motivated by a lack of theoretical relevance in foreigner talk research. Zuengler argues that most studies done on foreigner talk are descriptive in nature. For example, the linguistic features that characterize foreigner talk are well documented in the literature (See section 3.3). However, not all NSs use foreigner talk with NNSs, nor foreigner talk when used includes all the features listed in the literature. The literature lacks an explanation for the variation in foreigner talk use. Another well-documented phenomenon in foreigner talk use is variation in grammaticality. The absence of a dominant theory to explain foreigner talk use is prevalent in the fact that researchers’ dispute has been over whether to describe foreigner talk as grammatical or ungrammatical rather than over explaining why such variation in grammaticality occurs. Hence, Zuengler calls researchers to shift foreigner talk beyond a descriptive level and view it within a CAT framework to provide clarifications for its use and bring consistency to the literature.

3.4.2.1 Accommodation in NS-NNS interactions

Four dimensions from CAT presented in Coupland *et al.* (1988) are adopted to explain the dynamics and strategies underlying NS-NNS interactions: NS interactional goals, NS perceptions of the NNS, NS encoded strategies and NNS decoding of NS strategies (Zuengler 1991). The first three dimensions are the most relevant to the current discussion because the utmost interest of the current study is NS's speech in interaction with NNSs and not NNS's perception or interpretation of NSs strategies. Thus, further discussion of the first three dimensions will be provided in the coming paragraphs. It is important to note that Zuengler relies on foreigner talk studies from the 1970s and 1980s in his discussion.

According to CAT, interaction involves certain goals that a speaker wishes to achieve from communication; for example, communicating effectively, enhancing comprehensibility, gaining social approval and emphasizing distinctiveness from the listener (Coupland 1988; Giles and Gasiorek 2013). A number of studies have described foreigner talk as reflecting some of these goals, especially the NS's desire for successful communication and mutual comprehensibility (e.g. Taron 1980; Clyne 1981; Freed 1981; Snow *et al.* 1981; Hatch 1983). Zuengler states that the type of communicative goal the NS wishes to achieve depends on the type of interaction. For example, researchers (e.g. Long 1981; Hatch 1983; Ellis 1985) note more foreigner talk use and probably more concern for comprehensibility in a two-way interaction between NS and NNS rather than in a one-way interaction led by the NS. Furthermore, NSs aim more for increased comprehensibility when the interaction is spontaneous and when the topics are abstract and complex (e.g. Snow *et al.* 1981; Long 1983). Other than achieving mutual comprehensibility, Evans (1987) and Clyne (1980) reveal that NS's in their studies used foreigner talk to identify with the NNSs. This, according to Zuengler, reflects the second goal of communication which is gaining the interlocutor's social approval. The goal of emphasizing distinctiveness could be manifested in the non-use of foreigner talk (Zuengler 1991). Some foreigner talk studies demonstrated that NS's used foreigner talk to enhance the social distance and achieve distinctiveness from their NNS interlocutors (e.g. Valdman 1981; Perdue 1984).

The second dimension of CAT concerns the perceptions a speaker has of their interlocutor which can be based on the interlocutor's behavior or some stereotypes about the interlocutor (Coupland *et al*. 1988). Zuengler argues that NSs' perception of NNSs is based on two main dimensions: the NNS' ethnicity and the NNS' linguistic competence. Based on this claim, ethnic and cultural differences might either cause the NS to maintain distinctiveness by not using foreigner talk or might lead the NS to downplay the social distance and use foreigner talk to eliminate distinctiveness and achieve social approval. NNSs, however, might not be viewed as belonging to a specific ethnic group but rather as "foreign" (Harder 1980; Janicki 1986). According to Beeb and Giles (1984), NNSs in interaction with NSs can be viewed as a low social status because of their foreignness. The interactional goal of the NS can also change based on the perceived linguistic competence of the NNS. Some studies have demonstrated that the use of foreigner talk occurs with NNSs who are perceived to be of low communicative competence (Scarcella and Higa 1982; Long 1983; Ellis 1985).

The third dimension of CAT reflects the strategies the NS will use having one or more of the communicative goals discussed above and some perceived characteristics of the NNS interlocutor. These strategies are "convergence, divergence, maintenance and complementarity" (Zuengler 1991: 237). *Convergence* in foreigner talk reflects adaptations made in speech rate, vocabulary, and pronunciation. It ranges from few adaptations to nearly exact imitation of the NNSs speech. Evidence for this comes from descriptions in the literature that foreigner talk adjustments occur at varying degrees but occur most likely in interactions with low proficiency NNSs (Kleifgen 1985; Ellis 1985). Also, some observations in the literature pointed to the fact that foreigner talk contained ungrammatical utterances by way of imitating the NNSs mistakes. Zuengler concludes that the grammaticality of foreigner talk varies based on the extent of convergence to the NNS's speech but depending on the latter's degree of grammaticality. *Divergence* and *maintenance* are strategies used by NSs when they attempt to maintain distinctiveness from NNSs. Evidence from the literature for these strategies comes from observations on the lack of foreigner talk in NS-NNS interactions. NSs could also employ divergence by using a dialect or a jargon that the NNS does not use.

Complementarity is used as a strategy in foreigner talk when the NS attempts to accentuate role differences between interactants. Such interactions can represent such roles as high-status-low status (Beebe and Giles 1984), with NNSs assigned the low status role or non-foreigner-foreigner (Janicki 1986) or teacher-student (Zuengler 1991). Employing this strategy to emphasize role differences may or may not lead to using foreigner talk (Zuengler 1991). For instance, a NS in teacher-student interaction may use slower speech rate and clearer articulations in what they identify as a suitable style for a teacher. Likewise, a NS may emphasize the nonforeigner-foreigner role differences by using a standard variety if h/she perceives this to be the appropriate style for the foreigner (Harder 1980). Another NS who perceives themselves as a higher status may slow down their speech and use a more nonstandard speech (Zuengler 1987).

Studies that investigate the implications of CAT on NS-NNS interaction or FDS are scarce. One recent study that used CAT explicitly in explaining some of the FDS dynamics was that by Kangatharan (2014) described in Section 2.3.2. Kangatharan (2014) attributed vowel hyperarticulation in speech addressed to foreign sounding interlocutors partly to the strategy of convergence in which speakers accommodate to their interlocutors limited linguistic capacity.

The employer-FDH interaction setting relevant to the current study can indeed involve foreigner talk that is a function of different goals the NS is aiming to accomplish. Power relations and the way the NS perceives the FDHs in this context may or may not trigger FDS.

3.4.2 The Hyper and Hypo Theory

In speech communication, the hyper and hypo theory (H&H) is based on the assumption that speaking is a “goal-directed activity” (Lindblom 1996: 1687). Talkers attempt to select the linguistic elements of their utterances in a way that serves their communicative goal. This selection highly depends on the knowledge of the listener. As stated by Lindblom (1996), the speaker assesses the needs of the listener constantly and accordingly adapts his linguistic production (e.g. phonemes, words, syllables, etc.) based on those needs. This adaptation occurs along a continuum, with more powerfully articulated forms (hyper) at one end and less energetic articulations (hypo) at the other (Lindblom 1990). The hyper end of the continuum is characterized with a larger duration and amplitude of the different gestures as well as less overlap between gestures, whereas the hypo end of the continuum is characterized with a larger temporal overlap of gestures (Tabain 2002). As a further result, reduction and coarticulation are characteristic of hypospeech (Lindblom 1996). In the hypermode, on the other hand, the vowels and consonants are expected to reach their target form (ibid).

Lindblom (1996) also emphasizes that although one primary purpose in most speaking contexts is maintaining intelligibility, speech production serves several other functions that depend on the social context in which the speaker-listener conversation takes place. In fact, the speaker’s voice quality and pronunciation style are determined by communicative, physiological, cognitive and social factors (Fo´nagy 1983). Based on this, Lindblom (1996: 1688) further explains that

> Viewing speech processes in this manner, that is, in their broadest possible sociolinguistic and ‘‘ecological’’ setting, one cannot fail to be struck by the plasticity of human speech. Clearly, it is uncontroversial to suggest that, at a ‘‘macro’’ level, speech production is indeed adaptively organized and that the function of that adaptation is to successfully transmit the speaker’s communicative intentions to the listener.

Other than its relevance to the acoustic component of the present study, the H&H theory can also provide explanations to any variability in FDS related to the FDH context the current study is investigating. Given that most conversation that takes place between NS employers and FDH revolve mostly around home chores and making commands, speech adjustments can be influenced by the social context wrapped around these interactions. Thus, maintaining intelligibility may be one function of such interactions but also other functions that are specific to the employer-FDH context. Whether adjustments of speech production occur at the segmental level and help maintain intelligibility, Lindblom (1996: 1688) refers to ‘clear speech’ as being “informative” in this regard (Section 3.2).

3.5 Conclusion

This Chapter highlighted the linguistic characteristics of speech directed to different audiences, including infants, hearing-impaired listeners and foreigners. It described general features of each speech register but focused particularly on phonological and acoustic adaptations. The Chapter also presented two theoretical models that might be useful in explaining the dynamics of FDH-directed speech namely the H&H theory and CAT. Although these models differ in their focus, they agree that modifications in speech to NNSs are done to ease communication and achieve intelligibility.

The coming Chapter will focus on the linguistic variables and segments of interest to the current study. It will also discuss the sound system inventories of the target language, Nizwa Arabic, and that of the L1s of the foreign domestic helper participants. This will set the scene for discussing the methodology of the study (Chapter 5).

Chapter 4. The Linguistic Variables and the Sound Systems under Investigation

4.0 Introduction

In order to understand the nature of the phonological and acoustic input available to foreign domestic helpers (FDH), it is of paramount importance to provide a description of the sound system and syllable shape of the target language, Nizwa Arabic, as well as the phonemes and syllable shapes of interest to the present study. Additionally, understanding L2 learners' perception and production abilities requires an orientation to the L1 sound systems of the participants. Moreover, it may be hard to understand native speakers' (NS) motivation to adjust the phonological and acoustic properties of their speech when addressing L2 learners without having evidence of the latter's' limitations in the target language.

Hence, this chapter will first provide a description of the consonants of Arabic followed by a discussion of the consonants of interest to the current study and the rationale for choosing them. It will then provide a description of the acoustic correlates of fricatives and stops. This will be followed by a literature review on the articulatory and acoustic characteristics of the Arabic consonants of interest to the current study. After this, a discussion of Arabic vowels and the vowels of interest to the current study will be provided. This will be followed by a literature review on the articulatory and acoustic correlates of Arabic vowels. Then a discussion of the phonological aspects of Arabic consonant clusters and the consonant clusters the study examined will be presented. This will be followed by a description of the sound inventories of the FDHs' L1s. Finally, a comparison between the sound inventory of Nizwa Arabic and that of the L1s of the FDHs will be provided before the Chapter concludes.

4.1 Arabic Consonants

This section will present an overview of the consonants of the Nizwa dialect of Omani Arabic, the target language, as well as of Modern Standard Arabic (MSA) for comparison purposes. MSA is a modern version of Classical Arabic and is the official language of Arab countries (Fistcher 2013; Al Suwaiyan 2018).[8] It is also used as a medium of communication among the Islamic countries around the world (Fistcher 2013). Some differences between MSA and Classical Arabic exist in vocabulary and syntax (Al Suwaiyan 2018). However, no differences exist between the phonemes of MSA and

[8] Classical Arabic is the language of the Qur'an and classical literature (Al Suwaiyan 2018).

Classical Arabic (ibid.). Colloquial Arabic varieties that are used in every day communication also exist in Arabic speaking countries along with MSA. They differ from MSA in terms of syntax, vocabulary, morphology and phonology. Nizwa Arabic is a colloquial variety spoken in Nizwa, a city located in the interior suburbs of the Sultanate of Oman. The only available account of the sound system of Nizwa Arabic (Hamid *et al.* 2008) provides a phonological description of the dialect's vowels, consonants and syllable structure.

There are two main reasons why MSA is relevant to the current thesis. First, due to the absence of acoustic descriptions of the phonemes of Nizwa Arabic and due to major similarities between the phonemic inventories of MSA and Nizwa Arabic as will be shown below, I will rely on acoustic descriptions on MSA and other Arabic vernaculars in my literature review of the acoustic properties of the consonants under examination (Section 4.1.2). Second, some FDH participants had been exposed to Classical Arabic and/or MSA via Islam (Chapter 5). Hence, since some colloquial Arabic varieties differ from MSA/Classical Arabic in their phonology, it is essential to show what consonants Nizwa Arabic and MSA have in common.

MSA contains 29 consonants illustrated in Table 4.1. Nizwa Arabic has the full set of consonants found in MSA except for the voiced emphatic stop /dˤ/ which is realized as a voiced emphatic fricative /ðˤ/ in Nizwa Arabic and the voiced alveolar fricative /ʒ/ which is realized as a voiced palatal or velar stop /ɟ/, /g/ in Nizwa Arabic (Table 4.2). Hamid *et al.* state that both /dˤ/ and /ðˤ/ of MSA merge into one phoneme /ðˤ/ in Nizwa Arabic.

	Voicing	labial	Labio-dental	Dental	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	Voiced	b			d		ʒ			
	Voiceless				t		k	q		ʔ
Emphatic Plosive	voiced				dˤ					
	voiceless				tˤ					
Fricative	Voiced			ð	z			ʁ	ʕ	
	Voiceless		f	θ	s	ʃ		χ	ħ	h
Emphatic Fricative	Voiced			ðˤ						
	Voiceless				sˤ					
Nasal	Voiced	m			n					
Liquid	Voiced				ɾ l					
Emphatic Lateral	voiced				lˤ					
Approximant	Voiced	w				j				

Table 4.1 Consonants in MSA (adapted from Al-Ani 1970)

	Voicing	labial	Labio-dental	Dental	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	Voiced	b			d	(ɟ)	g			
	Voiceless				t	(c)	k	q		ʔ
Emphatic stop	voiceless				tˤ					
Fricative	Voiced			ð	z			ʁ	ʕ	
	Voiceless		f	θ	s	ʃ		χ	ħ	h
Emphatic fricative	Voiced			ðˤ						
	Voiceless				sˤ					
Nasal	Voiced	m			n					
Liquid	Voiced				ɾ l					
Approximant	Voiced	w				j				

Table 4.2 Consonants in Nizwa Arabic (Adapted from Hamid *et al.* 2008: 51). k and g are allophones of the palatal plosives according to the authors. They are realized as velars when they occur in emphatic environments or when they are next to a back vowel.

From the above tables, it is clear that MSA and Nizwa Arabic are quite similar in their phonemic inventories and differences between them are very minimal. The next section will provide a description of the consonants of interest to the current study.

4.1.2 The consonants under examination

In this study, I chose to examine the realization of a number of Arabic consonants by NSs when addressing FDHs as well as to examine the perception and production of these consonants by FDHs as L2 learners. Table 4.3 illustrates the consonants of interest to the current study using the realizations of Nizwa Arabic as the variables and their most common variants in Gulf Arabic[9] and Levantine[10] Arabic. Gulf Arabic and Levantine variants are reported because some of the FDH participants had been exposed to Gulf Arabic and/or Levantine (especially Lebanon) Arabic before moving to Oman (Chapter 5).

Variables (Nizwa Arabic)	tˤ	ðˤ	sˤ	χ	ʁ	q	ħ	ʕ	θ	ð	t	k	s
Gulf Arabic Variants	tˤ	ðˤ	sˤ	χ	ʁ	q/g/dʒ	ħ	ʕ	θ	ð	t	k/tʃ	s
Levantine Variants	tˤ/t	ðˤ	sˤ	χ	ʁ	q/ʔ/k/dʒ	ħ	ʕ	t/s	d/z	t	k/tʃ	s

Table 4.3 The linguistic variables and their variants in GA and Levantine Arabic (GA and Levantine variants were adapted from Biadsy *et al*. 2009: 55-56)

The consonants /tˤ, ðˤ, sˤ, χ, ʁ, q, ħ, ʕ, θ, ð / were included in the current study to examine their use by NSs when addressing FDHs, to examine FDHs' perceptual sensitivity to phonemic contrasts including these consonants and to examine the extent to which FDHs' production of these consonants is target-like. The first reason for choosing the emphatics / t ˤ, ðˤ, s ˤ/, the gutturals /χ, ʁ, q, ħ, ʕ/ and the interdental fricatives /θ, ð/ is their potential difficulty for L2 learners of Arabic. The emphatics involve a secondary articulation in addition to the primary one and thus could prove difficult for L2 learners of Arabic whose L1s do not contain similar features (Lababidi 2016). Al Mahmoud (2013) report that American learners of Arabic substitute the non-emphatic /t/ and /ð/ for their emphatic counterparts /tˤ/ and /ðˤ/. The gutturals involve places of articulation (uvular, pharyngeal, glottal) that are globally less frequent (marked), and which might be unfamiliar to many L2 learners of Arabic (Al-Mahmoud 2013; Hayes-Harb and Durham 2016). Al Mahmoud (2013) states that among the realization patterns American learners of Arabic used when learning Arabic was realizing the uvular

[9] Gulf Arabic is a name given to the language spoken in countries along the Southern East of the Arabian Peninsula including Saudi Arabia, Kuwait, UAE, Qatar, Bahrain and Oman (Biadsy *et al.* 2009).
[10] Levantine includes dialects spoken in Lebanon, Palestine, Jordan, Syria and Israel (Biadsy *et al*. 2009).

/q/ as [k] and the pharyngeal /ħ/ as [h]. The interdentals /θ, ð/ have been found to be difficult to acquire by both L1 and L2 learners (Major and Faudree 1996; Rau *et al.* 2009). In L1 acquisition, /θ, ð/ have been found to appear later than other consonants in children's speech (e.g. Wellman *et al.* 1931; Templin 1957; Ingram 1989; Clark 2009). Amayreh and Dyson (1998) found the same for children acquiring Arabic as an L1. In L2 acquisition, studies found that many L2 learners whose L1s lack /θ, ð/ pronounce these sounds using different variants. For example, Russian, Thai and Hungarian speakers of English realize the interdental fricative /θ/ as [t], whereas Japanese, Egyptian Arabic and German speakers of English realize /θ/ as [s] (Lombardi 2003). On the other hand, Cantonese native speakers of English from Hong Kong realize /θ/ as [f] (Peust 1996).

The consonant inventory of Gulf Pidgin provided by Smart (1990) gives evidence for the absence of the aforementioned consonants from the interlanguage system of foreign immigrants in the Gulf. Smart describes Gulf Pidgin as the language used between the indigenous Arabs and temporarily immigrant foreigners (Europeans and Indians) along the eastern Arabian coast from Kuwait to Oman. His description of Gulf Pidgin relied on personal observation as well as humorous material published in two Gulf newspapers. Smart reports that foreigners reduced the 29 consonants available in Gulf Arabic phonology to around 22 consonants when speaking Gulf Pidgin (Table 4.4).

	bilabial	labiodental	dental	alveolar	Alveo-palatal	palatal	velar	glottal
stops	P b		t d	s			k g	ʔ
fricatives		f v		z	(ʃ)	j	(x) (ɣ)	h
affricates					ʒ			
nasals	m			n				
tap				r				
approximants				l	j		w	

Table 4.4 The consonants of Gulf Pidgin Arabic (adapted from Smart (1990: 88). The consonants between brackets appeared in the orthography material and were in irregular use according to the author)

Furthermore, the description of the L1 inventories of the FDHS in Section 4.4 will provide further evidence for the absence of these consonants from the L1s of the FDHs. Thus, the present study was motivated to find out how NSs would use these particular consonants when addressing FDHs (e.g. whether they would accommodate to their FDHs pronunciation). Furthermore, the current study is interested in examining FDHs' perceptual discriminability of consonant contrasts that include these consonants as well as their accuracy in producing them.

The /t, s/ were included in the current study because they are the plain counterparts of the emphatics /t ˁ, sˁ/, and thus as we will see in Section 4.1.1, studies that investigate the emphatics always examine their non-emphatic counterparts to find out the acoustic correlates of emphasis. In addition, the present study is interested in finding out whether NSs would enhance the acoustic properties of the plain/emphatic consonants when addressing FDHs. The reason why the previous specific contrasts were chosen for the examination of consonantal enhancement is that their acoustic correlates have been examined thoroughly in the literature (Section 4.1.2.2). Additionally, these segments are less subject to variation among Arabic vernaculars compared to the other consonants. /t, s, k/ were also included as controls in the picture naming task used to elicit production data from FDHs.

4.1.2.1 Acoustic correlates of fricative and stop consonants

The present study examined any adaptations that exist in the NSs' production of the consonants described in the previous section. Auditory analysis was supported by a subsequent analysis based on spectrographic profiling of the acoustic patterns of the target consonants. Further acoustic analyses were carried out on the plain/emphatic alveolar fricatives /s/ and /sˁ/ as well as the plain /emphatic alveolar stops /t/ and /tˁ/. The linguistic variables that are to examine any acoustic adaptations in the plain/emphatic fricatives and stops are spectral moments (Centre of gravity, standard deviation, skewness, kurtosis), duration and intensity of the consonants as well as the formant frequencies of the vowels following them (F1, F2, F3). It is important to understand why these specific acoustic variables were chosen for examination in the current study. This section will provide a description of the most prominent correlates of fricatives and stops.

The main articulatory and acoustic features that characterize English fricatives are: (1) the formation of a narrow constriction some place in the vocal tract, (2) the formation of turbulence airflow, and (3) the maintenance of turbulence noise (Kent and Read 2002). Fricatives have long intervals of noise or aperiodic energy relative to stops and affricates, which characterize them as a sound class (ibid).

non-sibilant fricatives (e.g. /θ/, /ð/, /f/, /v/) have low-intensity noise in the higher frequencies of a spectrogram compared to sibilants (e.g. /s/, /ʃ/, /ʒ/, /z/) (Ladegfoged and Johnson 2001). This difference in noise between these classes of fricatives is essential to their perceptual distinction (Kent and Read 2002).

Previous studies of English fricatives have focused on four acoustic properties that can distinguish between fricative places of articulation and which will be helpful in distinguishing between the fricatives of interest here: spectral characteristics of the frication noise, amplitude, duration of the noise and cues of the transition from fricative to the adjacent vowel (Jongman *et al.* 2000). The spectral shape of fricatives is influenced by the size and shape of the oral cavity before the constriction. A more defined spectrum results from a longer cavity (Stevens 1998). Accordingly, labiodental and dental fricatives are characterized with a rather flat spectrum while alveolar and post alveolar fricatives display a well-defined spectral shape (Behrens and Blumstein 1988; Stevens 1960). The literature demonstrated that spectral properties of frication noise help to distinguish between the sibilant fricatives and the non-sibilant fricatives (Hughes and Halle 1956; Jongman 2000; Gordon *et al.* 2002). Additionally, thanks to spectral properties, the fricatives /z, s/ within the sibilant group can be distinguished from /ʒ, ʃ/ (Hughes and Halle 1956; Stevens 1960; Behrens and Blumstein 1988; Jongman 2000). Examining spectral moments is one of the techniques used to measure and quantify spectral properties and classify consonants. This technique allows for the extraction of a number of spectral features from distinct time segments in the speech signal including mean, variance, skewness and kurtosis (Nissen and Fox 2009). Jongman *et al.*'s (2000) account of spectral moments is briefly described here, as it provides one of the most straightforward descriptions. Measuring spectral moments relies on a statistical procedure used to classify obstruents, incorporating both local and global information of speech sounds. Local aspects include mean frequency while global aspects include spectral tilt and peakedness. Analysis of spectral moments may cover one or multiple regions of the consonant. Mean (also called center of gravity) represents the average energy concentration while variance (also called standard deviation) represents the range of energy concentration. Skewness refers to spectral tilt in phonetic terms. It reflects the overall slope of the energy distribution. Positive skewness indicates concertation of energy in lower frequencies and this suggests a negative tilt, whereas negative skewness indicates a concentration of energy in higher frequencies and suggests a positive tilt. Kurtosis reflects the peakedness of a distribution. High peakedness is obtained with high kurtosis values which also indicate a perceptibly defined spectrum while a flat distribution is obtained with low kurtosis values and indicates no clear peaks.

As for stops, a number of complex movements in the vocal tract are involved in their production. When stops are produced, pressure builds up behind a complete constriction in the vocal tract followed by a sudden release of the airflow resulting in a sudden increase in the energy, referred to as the burst (Fuchs and Birkholz 2019). Broadly, the acoustic correlates that have been suggested for the classification of stops fall into two types: properties based on the spectral features of stop burst and properties based on formant transitions from onset to a mid-position of the adjacent vowel (Prathosh *et al.* 2015). According to acoustic theory, properties that define the place of articulation of stop consonants can be obtained from examination of the short-time spectrum tested at consonant release (Fant 1960; Stevens and Blumstein 1978). The theory makes predictions about the properties of the consonant energy generated at the release and the estimated values of the likely frequencies of the vocal tract surrounding the consonant release (Stevens and Blumstein 1979). As a result of these formant locations and of the burst features, gross properties that are distinctive depending on the place of articulation can appear at a spectrum tested over 10-20 ms following the stop release. In the case of alveolar stops, the theoretically predicted property of the spectrum is diffuse energy throughout the frequency range, with a predominance of spectral energy at higher frequencies. Halle *et al.* (1957) state that bilabial stops have a concentration of energy in the lower frequencies at around 500-1500 Hz, whereas alveolar stops have a predominance of energy in a higher frequency range (above 4000 Hz). Velar stops have a concentration of energy in the intermediate frequency (1500-4000 Hz) range. When a stop is followed by a vowel, rapid changes in the formant frequencies (transitions) are also important cues for the perception of the different groups of stops (Halle *et al.* 1957). Using a more sophisticated statistical measure to classify word-initial voiceless obstruents based on spectral features, Forrest *et al.* (1988) confirms that mean, skewness and kurtosis of the burst spectra robustly distinguish different places for the articulation of stops. For example, they reported that /p/ and /t/ were distinct in skewness and mean, whereas velar stops like /k/ differed from other places of articulation in kurtosis but were similar to /p/ in mean and skewness. Another important cue for the identification of stop consonants is the transition in the formant frequencies of the following vowel (Halle *et al.* 1957). Other studies also examine the role of temporal cues in stop classification such as voice onset time (VOT) and closure duration (Prathosh *et al.* 2015). Lisker and Abramson (1964: 1) define VOT as "the time interval between the burst that marks release of the stop closure and the onset of quasi periodicity which reflects laryngeal vibration". VOT is a primary acoustic cue that distinguishes voiced from voiceless stops in many world languages (e.g. Lisker and Abramson 1964; Cho and Ladefoged 1999; Beckman *et al.* 2011). Additionally, VOT is one of the cues that can differentiate plain from emphatic stops in Arabic as will be discussed in Section 4.1.2.2. Thus, its examination is crucial for the current study.

The following section will focus primarily on the articulatory and acoustic features of the Arabic consonants of interest to the current study. This description will rely on research conducted on MSA and other Arabic varieties due the absence of acoustic studies on the consonants of Nizwa Arabic as indicated earlier.

4.1.2.2 The emphatics / tˤ, ðˤ, s ˤ/

Emphatic consonants are a class of coronal obstruents which are produced with a secondary articulation that involves a constriction in the upper pharynx and a primary articulation at the dental/alveolar area (Davis 1995; Zawaydeh 1998; Al Khatib 2008). In addition, other articulatory gestures play a role in emphasis production such as "retracted epiglottis, a raised larynx, a pressed/tense voice quality, and/or a protruded lip posture" (Al-Tamimi 2017: 1). The number and type of these coronal consonants differ with respect to the different Arabic verities. Most researchers agree that there are generally five emphatic consonants /tˤ, dˤ, sˤ, zˤ, ðˤ/ which contrast with their non-emphatic counterparts /t, d, s, z, ð/, some or all of which might be part of the different Arabic dialects' phonemic inventories (Zawaydeh 1998; Israel *et.al.* 2012; Algryani 2014). According to Khattab (2002), some researchers consider other consonants as emphatics such as the pharyngeals /ħ, ʕ/, and the uvulars /q, ʁ, χ/, whereas some Arabic varieties have established emphatic correlates of other consonants including /b, m, n, l, r/. Whichever opinion one chooses to adopt as to the range of emphatic consonants, this section takes the most common view and review studies that focus on the emphatics /tˤ, dˤ, sˤ, z ˤ, ðˤ/, provided that the current study will focus on the emphatics /tˤ, ðˤ, sˤ/ available in the dialect of the NS participants (Section 4.1.1).

Acoustic studies have been conducted to investigate the acoustic and perceptual correlates of emphatic consonants and their plain counterparts. Listeners can successfully discriminate between emphatics and their plain counterparts auditorily due to a difference in timbre between the two contrasts (emphatics have dark timbre while plain consonants have light timbre) reflected in changes in formant transitions and resonance frequencies (F1, F2, F3) (Khattab 2002). The most consistent acoustic correlate of emphasis that has been reported in different Arabic dialects is F2 lowering of the following vowel. Lowered F2 indicates retracted tongue among other things.

Al-Masri and Jongman (2003) investigate the acoustic correlates of emphasis in the northern dialect of Jordanian Arabic. They found a significant lowering of the F2 in the vowel adjacent to emphatic consonants compared to the same vowel in syllables with their plain counterparts (a drop of 521 Hz). Similarly, in a study investigating phonological and phonetic aspects of MSA emphatics and gutturals, Bin-Muqbil (2006) found that the most significant difference between emphatics and their plain counterparts is that F2 of vowels following emphatics was much lower than that following plain consonants. The results of another study by Jongman *et al.* (2007) who analysed data from eight male and female speakers of the Irbid dialect of Jordanian Arabic indicate that the vowel adjacent to the emphatic consonant was the most affected by emphasis and consistently showed lower F2 at the onset, midpoint and offset positions. Likewise, a recent study by Al-Tamimi (2017) demonstrates that F2 of vowels following the emphatic stop /dˤ/, in the productions of 20 Moroccan- and Jordanian-Arabic speakers, is significantly lowered compared to that following /d/. Hence, F2 has been found to be a highly reliable acoustic correlate of emphasis regardless of the dialects examined.

Changes in F1 of the following vowel could also be useful in providing information about the exact location of constriction in the upper pharynx (Jongman *et al.* 2011; Hassan and Esling 2007). For instance, a high F1 of the following vowel predicts a constriction at the post-uvular pharyngeal region whereas a low F1 indicates a constriction in the velar region (Hassan and Esling 2007). If, however, F1 is unchanged, this suggests a constriction in the uvular region (ibid.). It is important to note though that these changes in F1 should not be generalized in all environments as sometimes other articulators play a role in the production of emphasis such as lip protrusion and tongue sulcalization (Ali and Daniloff 1972).

Most studies examining F1 found that vowels adjacent to emphatic consonants show higher F1 compared to their plain counterparts. Giannini and Pettorino (1982) report that F1 and F2 of vowels in emphatic environments approached each other as F1 raised and F2 lowered. Yeou (1997) and Zawaydeh (1999) similarly found F1 to rise in emphatic environments. Jongman *et al.* (2007) and Jongman *et al.* (2011) report higher F1 of the vowel measured at onset, midpoint and offset positions in the Jordanian dialect. Similalry, Al-Tamimi (2017) reports higher F1 of vowels in the emphatic context. Higher F1 in these studies suggests a more open, pharyngeal constriction. On the contrary, early studies by Card (1983) and Norlin (1987) found that F1 was not consistently affected by emphasis in their study. The discrepancy in results of F1 patterning across these studies might be due to dialectal differences or methodological issues.

With regard to F3, findings on the location of secondary articulation and the coarticulatory influence of emphatics on following vowels have not been conclusive (Hassan and Esling 2007). For example, Giannini and Pettorino (1982), Card (1983) and Norlin (1987) found no consistent change in F3. However, Jongman *et al.* (2011) report higher F3 for vowels in emphatic environments compared to vowels in plain environments. Al-Tamimi (2017) also report higher F3 for vowels of Morrocan Arabic in an emphatic domain, though this was non-significant.

Research also shows that the degree of constriction caused by emphatics is quite variable depending on the manner of articulation and vowel environment. According to Laufer and Baer (1988) constrictions are less extreme for fricatives than for stops, and less extreme in a /i/ context than in a /a/ context. Similarly, Card (1983) and Yeou (1997) observe that F2 lowering at vowels midpoint for the vowels /a/, /i/ and /u/ differs, with the greatest lowering taking place for the vowel /a/, followed by /i/ and /u/. Jongman *et al.* (2007) and Jongman *et al.* (2011) report that F2 lowered the most in an /a/ environment while it lowered the least in an /u/ environment. Jongman *et al.* (2011) also observe that vowel length played a role in the degree of emphasis influence, with short vowels being the most affected by emphasis. In short vowels, F2 was significantly lower, F1 was significantly higher and F3 was lower compared to long vowels.

The effect of emphasis on the emphatic consonants themselves has also been reported in some studies. For instance, Jongman *et al.* (2011) report a considerably lower spectral mean (also termed 'centre of gravity') in stop emphatics compared to their plain counterparts. However, fricative emphatics did not show any effect of emphasis with regard to spectral mean. A similar result of the role of spectra is reported by Bin-Muqbil (2006) who found that the emphatic/plain continuant pairs were similar on their spectral moments whereas emphatic/plain stop pairs were significantly distinguishable from each other based on their spectral properties. Some studies report a difference in VOT between stop emphatics and their plain counterparts in some Arabic varieties. Several studies found that the voiceless alveolar stop /t/ was aspirated (it has longer VOT), while its emphatic counterpart /tˤ/ was unaspirated (it has shorter VOT) (e.g. Al-Ani 1970; Giannini and Pettorino; 1982 Al-Nuzaili 1993). Others found some dialectal or gender variability in the distinction between /t/ and /tˤ/ in terms of VOT (e.g. Heselwood 1996; Khattab 2002). For instance, Heselwood (1996) found that VOT for /t/ was longer than that for /tˤ/ in his four Iraqi speakers, which supports previous studies on the aspiration aspect. However, he found no pronounced difference between the same stop consonants in his Egyptian subjects who produced both as aspirated.

Thus far, we have seen that F2 is an important acoustic correlate of emphasis and is associated with physiological changes such as tongue backing, lip protrusion and tongue hollowing. F1 in an emphatic environment has been found to be raised in most studies. However, findings on the pattern of F3 have not been conclusive in the literature. As for the stop coronal consonant /tˤ/ specifically, lack of aspiration represented by a shorter VOT establishes an additional phonetic component that complements the others in some Arabic varieties. The spectral mean of stop emphatic consonants could also be a reliable acoustic correlate. Table 4.5 presents a comparison between emphatics and their plain counterparts with regard to the possible phonetic components associated to each.

	Articulatory	Acoustic
Emphatic	Pharyngeal narrowing Lip Protrusion Early glottal adduction	F1/F2 approximation F1, F2, F3 lowering Short VOT
Plain	Pharyngeal expansion Neutral lips Tongue front raising Later glottal adduction	F1/F2 separation F3 raising Longer VOT

Table 4.5 Comparison of acoustic and articulatory features associated with emphatic consonants and their plain counterparts (From Khattab 2002:144 and Al-Tamimi 2017:6)

4.1.2.3 The uvulars /χ/, /ʁ/ and /q/

Arabic has three uvular sounds: a voiceless continuant /χ/, a voiced continuant /ʁ/ and a voiceless unaspirated uvular stop /q/. Generally, uvulars are produced with the dorsum of the tongue retracted and raised towards the posterior wall of the oropharynx (McCarthy 1994). The three uvulars, however, differ slightly in their place of articulation and the degree of constriction. Based on x-ray data, Delattre (1971) found that during the production of /ʁ/, the uvula is curled downward towards the raised back of the tongue to produce a slight trill that can be hardly noticed. /χ/ is produced in a similar manner, but with a narrower constriction due to a slight raising of the larynx and negative participation of the uvula which is held flat at the back of the tongue. As with /q/, productions involve exaggerated tongue raising and retraction to achieve full constriction. Ghazeli's (1977) x-ray findings slightly differ from the previous account with regard to the tongue position. Ghazeli found that the tongue was retracted more during the production of /ʁ/ in her study. The location of constriction during the production of /χ/ was between that of /k/ and /q/.

Acoustically, the production of /χ/ features aperiodic noise in spectrograms according to both Al-Ani (1970) and Ghazeli (1977). Al-Ani's spectrograms point to a minimum spectral energy for /χ/ at around 1500 Hz before /i/, 1000 Hz before /a/ and 800 Hz before /u/. /ʁ/ is characterized to a certain degree with vowel–like formants throughout its duration. According to Ghazeli (1977), F1 for /ʁ/ ranges between 500 and 600 Hz next to /a/, while F2 ranges between 1200 and 1300 Hz. F1 is higher next to /a/ compared when next to /i/ or /u/.

Acoustic examinations also show that, similar to emphatics, uvulars spread emphasis into adjacent vowels. The domain and size of emphasis spread, however, have been reported to be different depending on the uvular sound and the following vowel. For example, according to Al-Ani (1970), F2 onset of /i/ following /χ/ and /ʁ/ is lowered to 1800-1900 Hz, while F2 of /u/ after the same sounds is raised to 1350 Hz. However, the coarticulatory effect from /ʁ/ on F2 onset of /a/ is stronger than that from /χ/. There is still much stronger articulatory effect from /q/ with F2 values for /i/, /a/ and /u/ reaching 1600 Hz, 1150-1200 Hz, and 900 Hz respectively. Bin-Muqbil's (2006) findings are similar to those of Al-Ani with respect to the low F2 transitions in neighboring vowels /a/ and /i/. However, contrary to Al-Ani, Bin-Muqbil found that F2 transitions in /u/ following uvulars was lower than when following emphatics. /χ/, /ʁ/ and /q/ were also found to affect not only F2 but also F1 of following vowels in a study by Giannini and Pettorino (1982). Emphasis spread from uvulars, however, have not been found to be as extensive as that from emphatics. There has been evidence that emphasis spread from uvulars does not reach adjacent consonants, adjacent high vowels or following syllables (e.g. Ghazeli 1977; Kuriyagawa 1984).

To sum up, Arabic uvulars /χ/, /ʁ/, /q/ are produced with two main movements: the dorsum of the tongue is pulled back and then raised towards the uvular region. F1 and F2 of the following vowels are important acoustic correlates for the identification of uvulars. Emphasis spread is evident with uvulars, but it is not as large as that for emphatics.

4.1.2.4 The pharyngeals /ħ/ and /ʕ/

Arabic pharyngeals /ħ/ and /ʕ/ are produced in the lower pharynx region, with two main articulators: the tongue root and the epiglottis (McCarthy 1994). The posterior pharyngeal wall and the tongue root move together inward in the production of pharyngeals. Also, during the production of pharyngeals, the degree of constriction in the pharynx is extremely high and less variable (Laufer and Baer 1988). This is opposite to what we have seen for emphatics, which show less extreme constriction and more variation in the degree of constriction. Delattre (1971) reports that the constriction for /ħ/ is lower and narrower than the constriction for /ʕ/.

Acoustically, pharyngeals have been characterized with vowel-like formant structures (Al-Ani 1970) and high F1 due to a lower place of constriction and wider oral cavity (Ghazeli 1977; McCarthy 1994). /ħ/ has been mostly described as a voiceless fricative or approximant (McCarthy 1994; Heselwood 2007). According to Ghazeli (1977), F1 for /ħ/ ranges between 550-1100 Hz. /ʕ/ has been most often classified as a voiced fricative, most probably due to pairing it phonologically with /ħ/ (Heselwood, 2007). F1 for /ʕ/ falls between 650-900 Hz (Ghazeli 1977). Spectrographic and more advanced analyses also show that F1 for pharyngeals is higher than F1 for uvulars (Alwan 1989). This is due to pharyngeals being produced with more posterior constriction compared to uvulars. Alwan (1989) examines the effect of a raised F1 in perceptual identification of pharyngeals using synthesized speech samples. The results indicate that, when F1 was raised, the sound was perceived as /ʕ/, whereas when F1 was lowered, the sound was perceived as a uvular /ʁ/. Alwan's conclusion highlights the importance of F1 as a perceptual cue for distinguishing sounds produced in the posterior region of the vocal tract and for distinguishing pharyngeals from uvulars. F2, however, was not found to have a significant effect on the identification of the voiced pharyngeals and the voiced uvular.

There has been some inconsistency in describing /ʕ/ due to dialectal, inter-speaker or methodological differences. For example, Al-Ani (1970) states that /ʕ/ was realized as a voiceless stop in the speech of his Iraqi speakers. On the contrary, Butcher and Ahmad (1987) found that /ʕ/ was realized as a pharyngeal approximant in the productions of 3 Iraqi speakers. Laufer (1996) studied the production of the pharyngeals /ʕ/ and /ħ/ by 23 Hebrew and Arabic speakers (dialect of the latter not provided). Acoustic analyses and physiological observations showed that in fluent speech /ʕ/ was realized as a pharyngeal approximant and never as a fricative. In this case, /ʕ/ did not show any frication and was rather detected due to raising of F1 and lowering of F2, which caused relatively wide constriction in the pharynx. Laufer's data also lends some support to Al-Ani's in that in some occasional cases (e.g. slow pronunciation), /ʕ/ was realized as a stop.

As to the coarticulatory effects of pharyngeals on neighboring vowels, the most consistent effect is a rise in F1. Al-Ani (1970) reports that F1 of the vowel /i/ is much higher when adjacent to word-initial /ʕ/ compared to its normal value (400 Hz or higher). F2 for /i/ is lower (1500 Hz) than its normal value. The F1 of vowels following /ħ/ is not mentioned in Al-Ani, but the F2 of the vowel /i/ lowers to 1750-1900 Hz. The vowels /a/ and /u/ are only slightly affected by /ħ/. Butcher and Ahmed (1987) similarly found that both /ħ/ and /ʕ/ are accompanied by a raised F1 at the steady state of following vowels.

In short, Arabic pharyngeals are produced with a raised larynx, mainly by approximating the tongue root to the posterior pharynx wall causing a constriction at the level of the epiglottis. They have been generally classified as fricatives, but they can be realized as approximants or stops (for /ʕ/) depending on speech rate or style. F1 of the pharyngeals as well as of the neighboring vowels is an important acoustic and perceptual cue, which can determine their distinctive features from other gutturals.

4.1.2.5 The interdentals /θ/ and /ð/

Arabic /ð/ and /θ/ have not occupied much of the attention of Arabic linguists as the gutturals have. Therefore, the literature on these consonants is very limited. The fricatives /ð/ and /θ/ are continuously reported as being interdental sounds, produced with the tip of the tongue positioned between the lower and upper front teeth (Bateson and Ryding 1967; Al-Ani 1970; Al-Karouri 1996). Al-Ani (1970) describes Arabic /ð/ as a voiced interdental fricative that can be detected acoustically from the weak formants that appear in spectrograms. The influence of /ð/ on the following vowels is evident in the lowering of F2 of /i/ and /iː/ to 1850-1900 Hz compared to their 2200 Hz F2 at steady state. F1 of these vowels after /ð/ is also affected as it raises slightly compared to when in steady state. F2 of /u/ and /uː/ after /ð/ is raised to 1500 Hz from 750-800 Hz. In the same account, Al-Ani describes Arabic /θ/ as a voiceless interdental fricative that can be detected acoustically from the random noise in spectrograms. F2 of /i/ and /iː/ lowers slightly after /θ/ while F2 of /u/ and /uː/ rises.

The previous background on the articulatory and acoustic characteristics of the Arabic consonants of interest will help reveal any variation in the realization of these consonants in speech addressed to FDHs. It will also help in explaining the acoustic pattern of NSs' production that will be presented as acoustic profiling later in Chapter 6. In addition, the acoustic description of the emphatic/non-emphatic consonants will help especially with establishing hypotheses regarding the acoustic properties that NSs would most likely use in case of phonetic enhancement.

4.2 Arabic Vowels

MSA has six vowels, three of which are short /i/, /a/, and /u/ and the other three are their long counterparts /iː/, /aː/, and /uː/ (Table 4.6). The duration feature of vowels in Arabic is phonemic (Al-Ani 1970). In addition to these vowels, Arabic dialects may have other vowels. Nizwa Arabic has two extra types of long vowels /eː/ and /oː/ and two diphthongs /aj/ and /aw/ (Hamid *et al.* 2008). The diphthongs always appear word-finally and might alternatively be described as a combination of a short vowel and a glide. Table 4.7 illustrates the vowels in Nizwa Arabic with examples.

	Front	Central	Back
High	i/iː		u/uː
Mid	e/eː		o/oː
Low		a/aː	

Table 4.6 The vowels of MSA

Vowel	Example	Glossary
/i/	/bint/	girl
/a/	/tamm/	confirmed
/u/	/ʕumr/	age
/iː/	/diːn/	religion
/aː/	/baːb/	door
/uː/	/ʕuːd/	branch
/eː/	/beːt/	house
/oː/	/joːm/	a day
/aj/	/ħaj/	alive
/aw/	/gaw/	air

Table 4.7 The vowels of Nizwa Arabic with examples (From Hamid et al. 2008: 52)

4.2.1 The vowels under examination

In addition to the consonants in Section 4.1.2, I chose to study the production of the three long Arabic vowels /aː/, /iː/ and /uː/ by NSs when addressing FDHs compared to when addressing the NS control. The reason for choosing these three vowels is due to previous research on clear speech, IDS and FDS that have examined the size of the acoustic vowel space covering the triangular area of three-point vowels /i, a, u/ (e.g. Khul *et al.* 1997; Liu *et al.* 2003) or /i, a, o/ (e.g. Bradlow *et al.* 1996). To recall from Chapter 3, this was done to assess the expansion of the vowel space area that is assumed to serve intelligibility or language development functions. To understand the origin of the link between vowel space area and vowel clarity, it is crucial to provide a brief overview of vowel perception and production.

Vowel perception involves two processes: categorization and identification (Rosner and Pickering 1994). Vowel categorization refers to how listeners manage to detect the different vowels in a language, whereas vowel identification refers to how listeners manage to identify the same vowel under different circumstances (e.g. by different speakers or in different contexts). Vowel categorization or perceived vowel quality has been assessed in many studies following the measure initially developed by Peterson and Barney (1952), which used the values of the first two formants (F1, F2) extracted from the steady-state portion (midpoint) of the vowel.[11] Kent and Read (2002:106) termed this classic view that describes vowels and their perception the 'simple target model'. This model proposes that vowels exist in a canonical form that is static across phonetic contexts and is adequately defined by the frequency values of F1 and F2 at the middle point of the vowel. In this model, vowels are represented acoustically in a two-dimensional space (Ladefoged and Johnson 2001). The vowel chart includes values of F2 plotted against values of F1. This vowel plot correlates roughly with the place of articulation (Rosner and Pickering 1994; Ladefoged and Johnson 2001). In articulatory terms, F1 correlates with tongue advancement, while F2 correlates with tongue height. Figure 4.1 shows the F1/F2 acoustic plot (based on Joos 1948) and the articulatory chart (based on Jones 1960). From the chart, we can see that for the vowel /i/ for example, F2 is high (tongue is advanced) and F1 is low (tongue is high). F1 is high in the production of /a/ indicating tongue lowering, whereas the value of F2 indicates an intermediate position for the tongue. /u/, on the other hand, is produced with the tongue lowered and retracted resulting in low values for F2 and F1. Hence,

[11] This is the part of the vowel following the influence of the preceding segment and prior to the influence of the following segment

many studies use the F1/F2 acoustic chart and measure vowel space area to assess speech intelligibility and clarity (e.g. Bradlow *et al.* 1996; Kuhl *et al.* 1997; Uther *et al.* 2007). Expanding the vowel space size suggests that vowels are produced in a way that makes them more distinct from one another which presumably improves the listener's identification and categorization ability.

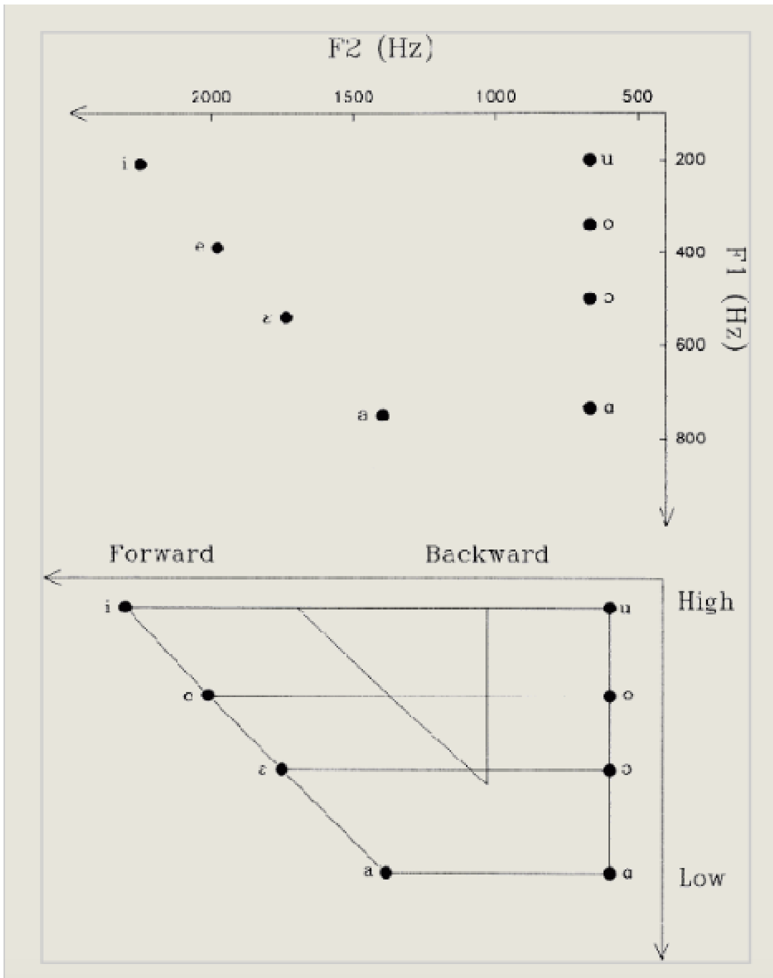


Figure 4.1 The F1/F2 acoustic chart and the articulatory chart (adopted from Rosner and Pickering 1994:47)

Although the classic vowel chart is a useful way to represent the vowels of a language (Ladefoged and Johnson 2001), it is oversimplified and has a number of limitations that pertain to vowel inherent properties and interspeaker differences (Kent and Read 2002). Vowels differ from each other not only in formant frequencies but also in other acoustic measures. For example, the traditional F1/F2 chart neglects vowel duration which is used as a perceptual cue in many languages (ibid.). Lehiste and Peterson (1961) found that tense or long vowels have larger durations than lax or short vowels. Also, they found that vowels differ in duration depending on the jaw position (vowels produced with open jaw are longer than vowels produced with a close jaw). Finally, they came to the conclusion that vowels differ in duration when produced with other sounds in context. Other factors that influence vowel duration include a set of suprasegmental properties (i.e. phonetic context) such as syllable stress, speaking rate and adjacent consonants (Kent and Read 2002). Experiments established that even though duration is not adequate in itself for the identification of vowels, it helps to improve the listener ability to recognize similar vowels such as the vowel set /ɑ/-/ɔ/-/ʌ/ (e.g. Hillenbrand *et al.* 2000). Another issue that the simple target model does not address is that vowels have temporal and dynamic differences and that vowel identification is not only achieved by the static state segments (Kent and Read 2002). One experiment demonstrates that vowels were identified accurately by listeners when only their transitions were presented in context (Jenkins *et al.* 1983).

Vowels also differ in their fundamental frequency of phonation ($f0$). Generally, vowel $f0$ changes with vowel height; high vowels have somewhat higher fundamental frequency than low vowels (Kent and Read 2002). When factors such as stress and intonation are controlled for, variation in inherent $f0$ can be detected (ibid.). However, Kent and Read (2002: 129) propose that variations in $f0$ might not reliably play a crucial role in vowel identification and that they may be used as 'secondary cues' or may be more salient when exaggerated by some speakers.

Variation in vowel acoustic properties is also a result of speaker characteristics. For example, differences in vocal-tract size could cause variation in formant frequency values. The shorter the vocal tract the higher the formant center frequencies and $f0$ values (Rosner and Pickering 1994). Peterson and Barney (1952) examine the distribution of American vowels produced by 77 speakers (children, males and females). They found that the acoustic distribution of vowels was varied across speakers. This is due to variability in the vocal tract size of these speakers. Females, for instance, have formant frequencies that are on average 15% higher than males' due to their shorter vocal tract (Al-Tamimi and Barkat-Defradas 2003). Children have the highest formant frequencies and F0 as their vocal tract size is the shortest compared to males and females (Rosner and Pickering 1994).

To account for the limitations of the classic target vowel model, some researchers have suggested elaborated models. Many of these models account mainly for the issue of speaker normalization.

According to Johnson (2005: 1), speaker normalization is based on the fact that “phonologically identical utterances show a great deal of acoustic variation across talkers, and that listeners are able to recognize words spoken by different talkers despite this variation”. Thus, this model provides an explanation for how listeners can recognize vowels when the primary acoustic cues for their distinctiveness (F1/F2) are ambiguous. A model that takes into consideration speaker normalization typically transforms the acoustic measurements of formant frequencies into a psychophysical space scaled in mels or Barks (Kent and Read 2002). This transform is assumed to be similar to that applied by the auditory system and thus can solve the speaker normalization issue.

Another model that found both the simple target model and the elaborated target model limited in accounting for vowel perception is the ‘dynamic specification model’ by Strange (1987). In this model, vowel identity is defined not only in terms of information obtained from their steady state portion but also based on information obtained from their formant transitions into and out of the steady state as well as the duration of the steady state.

The current study will combine approaches from the classic target model as well as the other models to report on any variation in NSs’ realization of the corner vowels when addressing FDHs. It will use the classic target model to explore any variation on vowel space. It will also use speaker normalization to account for any interspeaker variation and will also examine vowel duration, *f*0 and intensity following previous research (Uther *et al.* 2007; Knoll *et al*. 2015).

4.2.1.1 The Arabic corner vowels /iː/, /uː/ and /aː/

Since the current study is interested in the examination of the Arabic long vowels, the focus of the following description of Arabic vowels will mainly cover long vowels.

Al-Ani (1970: 23-24) describes Arabic /iː/ as a long high front unrounded vowel, /uː/ as a long high back rounded vowel and /aː/ as a low long central unrounded vowel. These vowels have a number of allophones that vary based on the phonetic environment. For instance, following pharyngealized (emphatic) consonants, vowels are backed which indicates lower F2 values (Ghazali 1977; Elgendi 2001). A number of studies also describe Arabic vowels acoustically by reporting on their F1/F2 values in isolation and in connected speech (e.g. Al-Ani 1970; Belkaid 1984; Abu Haidar 1994).

Figure 4.2 shows scatterplots of the F1/F2 values of Arabic short and long vowels in isolation as reported in different studies. Figure 4.3 shows F1/F2 values of the same vowels in connected speech.

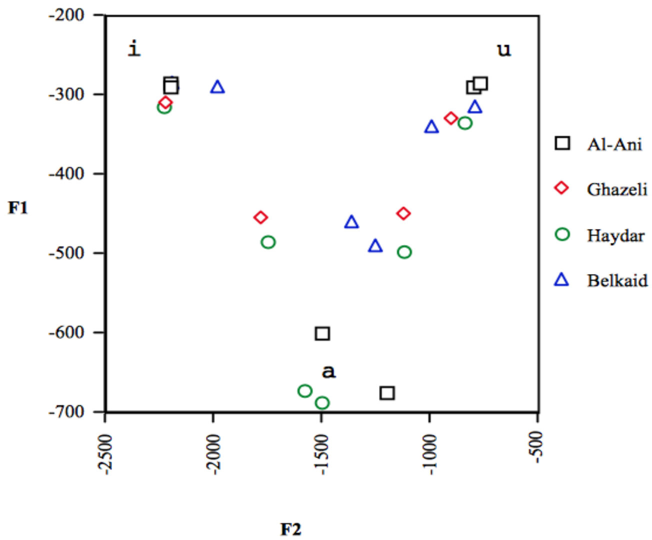


Figure 4.2 F1/F2 measurements for the six Arabic vowels in Isolation (From Newman 2002: 71)

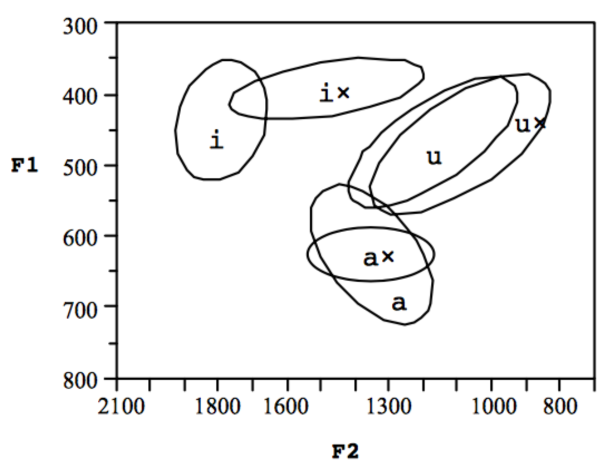


Figure 4.3 F1/F2 measurements of the six Arabic vowels in connected speech (From Newman 2002: 71)

A comparison of vowel properties among these studies is inappropriate due to some limitations pertaining to either flaw in their methodological frameworks, the small number of informants, different dialects and variants investigated or different phonetic environments (Saadah 2011). Providing details of vowel properties of different Arabic dialects as found by previous studies is also beyond the scope of this paper as the utmost interest of the current study is to examine whether the vowel acoustic space of NSs of Nizwa Arabic is expanded when addressing NNs as reported in IDS and FDS studies.

In short, we have seen that the target vowel model has been used widely in studies to represent the distribution of vowels in the acoustic space using F1/F2 values. This model has been extended to account for other phenomena that have been found to be useful in the identification and categorization of vowels in perception such as vowel inherent properties (e.g. duration, F0) as well as account for variation in the acoustic properties due to interspeaker differences (e.g. males vs. females).

4.3 Arabic Consonant Clusters

Final consonant clusters are very frequent in MSA. They do occur in some Arabic dialects (e.g. Cairo, Damiette, Tunisia, Maltese) though less frequently than MSA (Haddaad 1984; Kiparsky 2003). Optionally, an epenthetic vowel can be inserted to break up these clusters depending on speech style and dialect as in (4.1).

(4.1)

a. /Kata**b-t**/ ⟶ /kata**bɪt**/ ‘I wrote’

(Adapted from Kiparsky 2003)

Initial consonant clusters are not permissible in MSA (Haddaad 1984). However, phrase-initial consonant clusters are frequent in some Arabic spoken vernaculars (e.g. Iraq, Syria, Tunisia…etc) which allow the deletion of high short vowels in open syllables (a process called syncope) (Kiparsky 2003). For example, a word like /kɪtaːb/ ‘book’ is realized as /ktaab/ in these dialects. The cluster can also be broken up by inserting a prothetic vowel preceded by a glottal stop such as in /ʔɪktaab/. In addition to being in an open syllable, stress is another factor that conditions short vowel deletion in modern Arabic dialects (Watson 2011). For instance, Brame (1974) provided evidence that a short vowel in an unstressed syllable is deleted in Maltese.

According to Hamid *et al.* , Nizwa Arabic has five basic syllable types: CV, CVV, CVC, CVVC and CVCC. However, Nizwa Arabic also has a syllable structure of the type CCV (Ambu Saidi 2019) and CCVCC, CCCVCC (e.g./stroħt/).

Nizwa Arabic undergoes the same short vowel deletion in initial unstressed open syllables (Ambu Saidi 2019). However, According to Ambu Saidi (2019), all short vowels could be deleted in this environment and that unlike previous studies, vowel height seems to play no role in this process as shown in (4.2):

(4.2)

a. **bɪsˤ**aːtˤ → **bsˤ**aːtˤ 'mat'

b. **ma**nːtˤuq → **mn**ːtˤuq 'areas'

c. **nu**guːm → **ng**uːm 'stars'

(Adapted from Ambu Saidi 2019)

According to Ambu Saidi (2019), variation in the use of the resulting clusters exists among NSs of the dialect. NSs may or may not use syncope in their speech depending on external factors such as speech style (formal vs. informal), age and gender.

4.3.1 The consonant clusters under examination

The consonant clusters the study is interested in investigating are initial (onset) and final (coda) two-member consonant clusters. Given the two different facets of this study, the rationale behind choosing consonant clusters is that the production of these segments by NSs and foreigners may undergo variation for different reasons pertaining to each groups' linguistic situation. As discussed in Chapter 2, evidence exits as to the difficulty L2 learners face when acquiring a syllable structure different than that of the L1 or that does not conform to universal principles. Hence, it would be interesting to examine the production patterns of the FDHs in this study when producing the L2 consonant clusters and find out whether their production reaches native-like performance.

The study of consonant clusters is also interesting when it comes to how they are produced by NSs when addressing FDHs in comparison to the NS control. Considering that complex syllable structures are difficult to acquire by L2 learners (Eckman 1977; Broselow 1983) and that simplified input may facilitate language communication (Long 1983, 1985), it is possible that NSs of a language might adapt their complex syllable structure and simplify it in communication with NNSs.

Since NSs of Nizwa Arabic will either delete or retain the short vowel in initial unstressed open syllables, the current study is motivated to investigate whether addressing FDHs may trigger more retention of the short vowel for the purpose of producing a less complex syllable structure and converging to the simplified Arabic used among foreigners. No study, as far as I know, has examined coda consonant clusters in the Nizwa dialect of Omani Arabic. As a speaker of the dialect, it is safe to say that no variation exists in the production of coda consonant clusters in Nizwa Arabic and that NSs constantly maintain the cluster throughout their productions unlike what Kiparsky (2003) reported about some Arabic dialects that optionally insert a vowel to break up the cluster. However, there is still a possibility that NSs would break up coda clusters when addressing foreigners to simplify it. Thus, this study will examine the use of coda consonant clusters by NSs and FDHs in addition to onset consonant clusters.

4.4 The L1s Under Investigation

Figure 4.4 illustrates the migration route of the FDH participants in the current study from their countries of origin to Oman. The FDH participants migrated from nine countries in Asia and Africa: India, Bangladesh, Sri Lanka, Indonesia, the Philippines, Ethiopia, Uganda, Zanzibar and Nigeria. Therefore, they spoke a diverse set of L1s including: Telugu, Bengali, Sinhala, Indonesian, Tagalog, Oromo, Luganda, Swahili and Yoruba, respectively. A full description of the social and linguistic characteristics of FDHs will be provided in Chapter 5. It should be noted that for some L1s, recent linguistic accounts are not extensively available. Thus, the present study will rely on some old sources to describe these L1s.

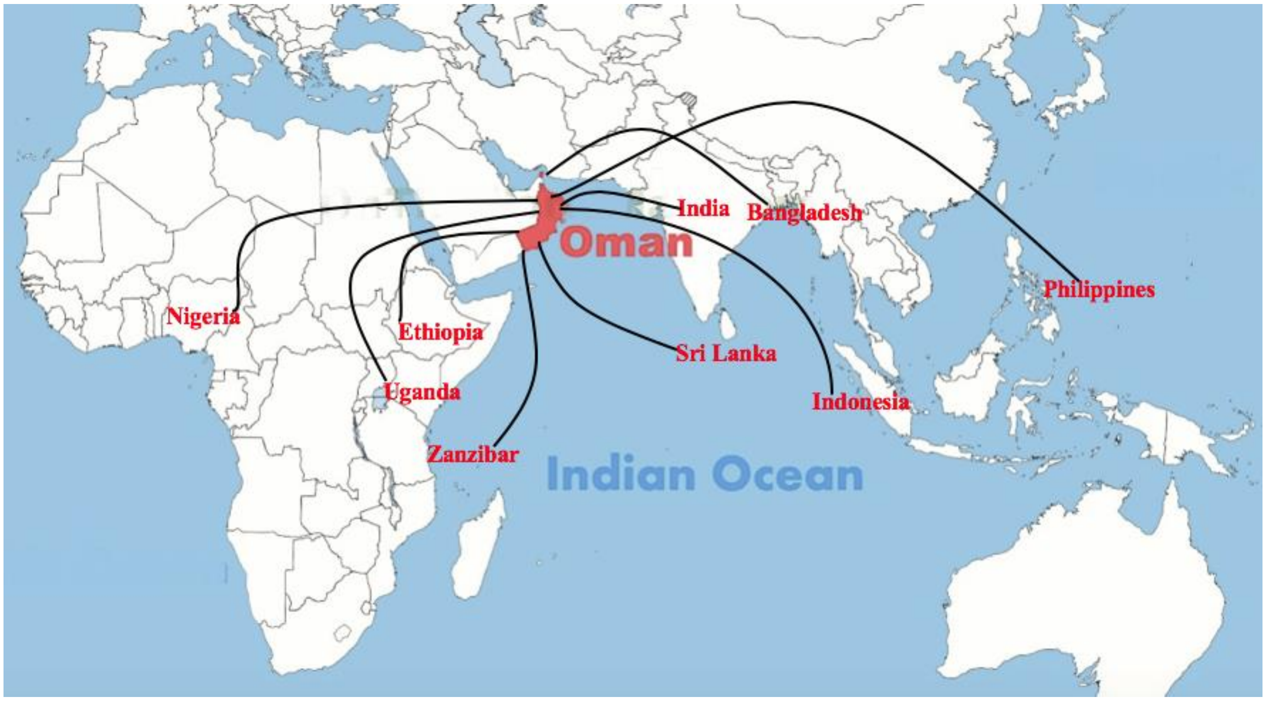


Figure 4.4 The migration path of FDHs from their countries of origin to Oman (the background map was adopted from www.google.image.com)

The following sections will provide a description of the sound inventories of the FDHs' L1s.

4.4.1 Telugu

Telugu is a Dravidian language spoken by around 7.19% of the Indian population (Bhaskararao and Ray 2017). Telugu vowels are shown in Table 4.8 All vowels contrast for length except /ɛ/ and /æː/ Bhaskararao and Ray 2017). Table 4.9 illustrates the consonants found in modern Telugu. This inventory includes phonemes that have been added to the native phonemic system due to the influence of different languages such as English, Sanskrit, Prakrit and Perso-Arabic over centuries (see Bhaskararao and Ray 2017 for more details of the borrowed sounds and the languages they came from). According to Bhaskararao and Ray (2017), the extended sound inventory reflects the speech of well-educated people. All consonants of interest to the present study are absent from the sound inventory of Telugu except for /t, s, k/ though these might also differ at the acoustic level. Thus, Telugu speakers are expected to face difficulty producing these consonants.

	Front	Central	Back
High	i/iː		u/uː
Mid	e/eː ɛ	ə/əː	o/oː
Low	æː	ɐ/ɐː	

Table 4.8 Vowels in Telugu (From Bhaskararao and Ray 2017)

	Voicing	labial	Denti-alveolar	Alveolar	Retroflex	Velar	Palato-alveolar	Palatal	Glottal
Stop	Voiced	b bʰ	d dʰ		ɖ ɖʰ	g gʰ			
	Voiceless	p pʰ	t tʰ		ʈ ʈʰ	k kʰ			
Stop Affricate	Voiced		dʑ				dʒ		
	Voiceless		tɕ				tʃ		
Fricative	Voiceless	f	s		ʂ		ʃ		h
Nasal	Voiced	m		n	ɳ				
Trill	Voiced			r					
Lateral	Voiced			l	ɭ				

Table 4.9 Consonants of Telugu (From Bhaskararao and Ray 2017)

Onset consonant clusters occur more frequently in borrowed words in Telugu, while coda consonant clusters are absent (Bhaskararao and Ray 2017). The only consonants that can occur word finally in the language are /m/, /w/ and /j/, otherwise all words end with a vowel (Sailaja 1999). The clusters of interest to the present study are absent from the syllable structure of Telugu. Thus, Telugu speakers are predicted to face difficulty producing onset and coda consonant clusters.

4.4.2 Bengali

Bengali is a world language spoken by over 175 million people in Bangladesh and Eastern India (ud Dowla Khan 2010). It is a descendent of Sanskrit and belongs to the Indo-Aryan subfamily of the Indo-European family (Ferguson and Chowdhury 1960; Thompson 2012). The vowels of Bengali are shown in Table 4.10. These vowels do not contrast in duration nor nasality (ud Dowla Khan

2010). The consonants in Bengali are shown in Table 4.11. The guttural, emphatic and dental consonants of interest to the present study are absent from the consonant inventory of Bengali. Hence, Bengali speakers are expected to face difficulty producing these consonants in the present study.

	Front	Central	Back
High	i		u
Mid	e ɛ		o ɔ
Low		a	

Table 4.10 Vowels in Bengali (From ud Dowla Khan 2010)

	Voicing	labial	Labio-dental	Dental	Post-alveolar	Velar	Glottal
Plosive	Voiced	b bʰ		d̪ d̪ʰ	d dʰ	g gʰ	
	Voiceless	p pʰ		t̪ t̪ʰ	t tʰ	k kʰ	
Stop	Voiced				dʑ dʑʰ		
Affricate	Voiceless				tɕ tɕʰ		
Fricative	Voiceless		f	s	ʃ		h
Affricate	Voiced						
	Voiceless						
Nasal	Voiced	m			n	ŋ	
Approximant	Voiced				ɹ		
Lateral	Voiced				l		

Table 4.11 4.11 The consonants of Bengali (From ud Dowla Khan 2010)

The syllable of Bengali consists of the following basic shapes: V, CV, VC, CVC, CCV, CCCV, CCVC (Rashel 2012; Thompson 2012). coda clusters are absent in Bengali, though some might occur in foreign words in the speech of multilingual Bengalis (Fergueson and Chowdhury 1960). Onset consonant clusters occur in standard colloquial Bengali and include the following combinations: stop+/l/ or /r/, /s/+stop or dental liquid and three-consonant clusters such as /str/ (Fergueson and

Chowdhury 1960; Thompson 2012). From this, Bengali speakers are predicted to face more difficulty producing Arabic coda consonants compared to onset consonants.

4.4.3 Sinhala

The Sinhala language belongs to the Indo-Aryan subfamily which is a member of the Indo-European family (Wasala and Gamage 2005). Being the official language of Sri Lanka, Sinhala is the mother tongue of around 74% of the population (ibid). Vowels in Sinhala are illustrated in Table 4.12. According to Smith (2001), all Sinhala vowels can be long and short phonemically except for /ə/ which rarely occurs long. It is also common for these vowels to occur in combination; available diphthongs include: /ei, ai, æi, ui, oi, eu, au, æu, ou, ae/.

	Front	Central	Back
High	i/iː		u/uː
Mid	e/eː	ə/əː	o/oː
Low	æ/æː	a/aː	

Table 4.12 Vowels in Sinhala (From Smith 2001)

Table 4.13 illustrates the 26 consonants including pre-nasalized sounds unique to the language that occur in Sinhala. The guttural, emphatic and dental consonants of interest to the present study are absent from the consonant inventory of Sinhala. Hence, Sinhala speakers are expected to face difficulty producing these consonants in the present study.

Local Sinhala consists of the following basic syllable structures: V, CV, VC, CVC (Wasala and Gamage 2005). As Sinhala has also loanwords and derived words from other languages, the syllable structures of these is mostly adapted to adhere to the syllable structure of local Sinhala (ibid). From this, Sinhala speakers are predicted to face difficulty producing onset and coda consonant clusters in the present study.

	Voicing	labial	Dental	Alveolar	Retroflex	Velar	Palatal	Glottal
Stop	Voiced	b	d		ɖ	g		
	Voiceless	p	t		ʈ	k		ʔ
Fricative	voiceless	f	s					h
Affricate	Voiced					ɟ		
	Voiceless					ç		
Pre-nasalized stop	Voiced	b̃	d̃		ɖ̃	g̃		
Nasal	Voiced	m		n			ŋ	
Trill	Voiced			r			ɲ	
Lateral	Voiced			l				
Glide	Voiced	ʋ					j	

Table 4.13 4.13 The consonants of Sinhala (From Smith 2001)

4.4.4 Indonesian

Indonesian, the official language of Indonesia, is an Austronesian language spoken by a large number of people as a first or a second language (Soderberg and Olson 2008). There are five main vowels in Indonesian shown in Table 4.14. Three diphthongs (/aj, aw, oj/) also occur in the language but only word-finally. Table 4.15 shows the consonants that occur in Indonesian. The consonants in parenthesis appear only in loan words and may show variation in their pronunciation (Soderberg and Olson 2008). The consonants of interest to the present study are absent from the consonant inventory of Indonesian. Accordingly, Indonesian speakers are expected to face difficulty producing these consonants.

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Table 4.14 Vowels in Indonesian (From Soderberg and Olson 2008)

	Voicing	labial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Stop	Voiced	b			d			g	
	Voiceless	p		t̪				k	(ʔ)
Affricate	Voiced					dʒ			
	Voiceless					tʃ			
Fricative	Voiced				(z)				
	Voiceless		(f)		s	(ʃ)			h
Nasal	Voiced	m		n			ɲ	ŋ	
Flap/Trill	Voiced				r				
Lateral	Voiced				l				
Approximant	Voiced	w					j		

Table 4.15 The consonants of Indonesian (From Soderberg and Olson 2008)

Indonesian has a simple syllable structure involving two basic shapes: CV and CVC (Suyanto *et al.* 2016). Open syllables are more common in Indonesian than closed syllables (ibid). Hence, Indonesian speakers are predicted to face difficulty producing Arabic onset and coda consonant clusters.

4.4.5 Tagalog

Tagalog (officially named Filipino) is one of two official languages besides English spoken in the Philippines (Schachter and Reid 2008). It belongs to the Austronesian language family and is the native language of Manila, the largest city in the Philippines (ibid). Tagalog phonology has been considerably influenced by the many loanwords from English and Spanish that have been incorporated into the language. Five vowels exist in contemporary Tagalog shown in Table 4.16 (Schachter and Reid 2008). Tagalog consists of 17 consonants illustrated in Table 4.17. Several other phonemes that have been adopted from loanwords can also be heard in the speech of many educated Tagalog speakers such as the labiodentals /f/ and /v/, and the affricates /tʃ/ and /dʒ/ (French 1988).

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Table 4.16 Vowels in Tagalog (From Schachter and Reid 2008)

	Voicing	labial	Dental	Alveolar	Velar	Palatal	Glottal
Stop	Voiced	b	d		g		
	Voiceless	p	t		k		ʔ
Fricative	Voiced			s			
	Voiceless						h
Nasal	Voiced	m	n		ŋ		
Lateral	Voiced			r, l			
Glide	Voiced	w				j	

Table 4.17 The consonants of Tagalog (From Schachter and Reid 2008)

Llamzon (1966) classifies the syllable structure in Tagalog into open syllables that contain structures such as V, CV, CCV, S-V, where S is a semi vowel and closed syllables that contain structures such as VC, CVC, CCVC, CVCC, CCVCC, VS-V, CVS-V, CV-VVC. Schachter and Reid (2008) highlight that the most common syllable patterns are CV and CVC. They also explain that in native Tagalog, onset consonant clusters are restricted to a consonant and a glide (e.g. /bwan/ ‘month’). Final consonant clusters are only available in loanwords from English. Thus, Tagalog speakers might be able to produce some Arabic onset consonant clusters but they will face difficulty producing coda consonant clusters.

4.4.6 Oromo

Oromo is a Cushitic language of the Afroasitic family spoken by over seven million speakers in and from Ethiopia and Kenya (Lloret 1995; Lloret 1997). Oromo phonemically distinguishes between five long and five short vowels shown in Table 4.18. The consonants of Oromo are shown in Table 4.19. According to Lloret (1995), the fricative velar [x] appears in Eastern Oromo instead of /k/ while Southern Oromo has [x] as an allophone of /k/. Oromo speakers are expected to face difficulty producing the consonants of interest in the present study due to the absence of these consonants from the Oromo inventory.

	Front	Central	Back
High	i/iː		u/uː
Mid	e/eː		o/oː
Low		a/aː	

Table 4.18 Vowels of Oromo (From Lloret 1997)

	Voicing	labial	Labio-dental	Denti-alveola	Palatal	Velar	Glottal
Stop	Voiced	b		d ɗ	dʒ	g	
	Voiceless	p’		t t’	tʃ tʃ’	k k’	ʔ
Fricative	Voiced						
	Voiceless		f	s	ʃ	(x)	h
Nasal	Voiced	m		n			
liquid	Voiced			r l	ɲ		
Approximant	Voiced	w			j		

Table 4.19 Consonants of Oromo (From Lloret 1995)

With regard to consonant clusters, consonants in Oromo do not appear in clusters word initially but they do so sometimes word medially (Lloret 1997). Consonants are not permitted word finally in Oromo; hence, words always end with a vowel (ibid). Oromo speakers are, therefore, predicted to find it difficult to produce Arabic onset and coda consonant clusters due to the dissimalrity between the L1 and L2 syllable structures.

4.4.7 Luganda

Luganda is one language among several others spoken in Uganda (Ladefoged *et al.* 1972). It belongs to the Bantu language family but it is a tone language (Ladefoged *et al.* 1972; Katamba 1974). Luganda has five basic vowels shown in Table 4.20.

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Table 4.20 Vowels in Luganda (From Katamba 1974)

The Luganda consonantal system is illustrated in Table 4.21. This shows that the consonants of interest are absent from the Luganda sound system. Thus, Luganda speakers are expected to find it hard to pronounce marked Arabic consonants.

	Voicing	labial	Labio-dental	Alveolar	Palatal	Velar
Stop	Voiced	b		d	ɟ	g
	Voiceless	p		t	c	k
Fricative	Voiced		v	s		
	Voiceless		f			
Nasal	Voiced	m			ɲ	ŋ
Lateral	Voiced			l		
Glide	Voiced	w			j	

Table 4.21 The consonants in Luganda (From Katamba 1974)

Possible syllable types in Luganda include CV(V), CVC, VC (Hyman and Katamba 1999). Consonants syllable finally are always glides (ibid). Ladefoged *et al.* (1972) also report that in all Uganda Bantu languages, consonant combinations of the type consonant+glide (e.g /bj/ and /bw/) are permissible. Consonants can also appear preceded by nasals (e.g. /mba/, /mva/, /mga/) (Ladefoged *et al.* 1972). Uganda speakers will find coda consonant clusters hard to produce but might be able to produce some Arabic onset clusters.

4.4.8 Swahili

Swahili, the official language in Zanzibar, is spoken as a first language by around one million people and as a second language by others as well as being used as a lingua franca by many African tribes

(Polome 1967; Mohammed 2001). Like Luganda, Swahili belongs to the huge family of Bantu languages. It has borrowed words from other languages such as English, Arabic, German, Portuguese and other Bantu languages (Mohammed 2001). Foreign words incorporated into the language have become completely "Swahilized" in their phonological structure (ibid).

At some time when Zanzibar and parts of East Africa were under the colonization of the Sultanate of Oman, Arabic had expanded as the language of the court and of some ceremonial occasions in Zanzibar due to the upsurge of Arab nationalism (Polome 1967). However, ever since independence, Arabic has been completely excluded from public life and has only been used by Muslims for the performance of their spiritual duties (ibid). Swahili consists of five cardinal vowels (/i, e, u, a, o/ and other secondary vowels shown in Table 4.22.

	Front	Central	Back
High	i		u
	ɪ		ʊ
Mid	e		o
	ɛ		ɔ
Low		a	

Table 4.22 The vowels of Swahili (From Mohammed 2001)

According to Mohammed (2001), 26 consonants are found in Swahili illustrated in Table 4.23. Polome (1967) argues that the interdental fricatives /ð/, /θ/ and the velar fricative /ɣ/ only occur in the speech of educated Swahili speakers living in the islands and coastal areas. Speakers who are less influenced by Arabic, particularly those from the inland, tend to realize /ð/ as [z], /θ/ as [s] or [t], and /ɣ/ as [g]. Additionally, Polome (1967) claim that the velar fricative [x] is an allophone of /h/ and occurs only in words borrowed from Arabic. Further to that, speakers who are less influenced by Arabic are more likely to realize /x/ as [h] in Arabic loanwords. To this end, the fricative interdentals and velars in the Swahili phonemic inventory proposed by Mohammed (2001) seem to have been incorporated into the language from Arabic and other foreign languages and are not part of the original Swahili sound inventory.

	Voicing	labial	Labio-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stop	Voiced	b			d	dʒ		g	
	Voiceless	p			t	tʃ		k	
Stop Affricate	Voiced						dʑ		
	Voiceless								
Fricative	Voiced		v	(ð)	z			(ɣ)	
	Voiceless		f	(θ)	s	ʃ		(x)	h
Nasal	Voiced	m			n	ɲ			
Liquid	Voiced				r l				
Approximant	Voiced	w					j		

Table 4.23 The consonants in Swahili (From Mohammed 2001). The consonants between brackets appear in Arabic words

Swahili has two permissible basic syllable patterns: V and CV (Batibo 2002). Initial consonant clusters are infrequent in Swahili and those that occur in the language come from foreign words that have been incorporated into the language such as *skati* from English *skirt* (Mohammed 2001). There are also initial consonant clusters that are common among Bantu languages that consist of nasal+ consonant (e.g. /mv/, /nd/, /ng/). However, the treatment of these combinations as a cluster is still debatable due to their behavior that resembles single segments in connected speech (ibid). Consonants word finally are not permissible in Swahili, though, some consonants may occur word finally in loanwords (Polome 1967; Mohammed 2001; Batibo 2002). To this end, Swahili speakers will find coda consonant clusters hard to produce but might be able to produce some Arabic onset clusters.

4.4.9 Yoruba

Yoruba is a language spoken in Southwest Nigeria by several million people (Owolabi 2012). It belongs to the Niger-Congo family, a mother of several other West African languages (Olmsted 1951; Owolabi 2012). Yoruba consists of seven basic vowel phonemes shown in Table 4.24. All vowels of Yoruba can occur long or short as well as oral or nasal (Olmsted 1951). Table 4.25 shows the consonants that occur in Yoruba's sound inventory. The marked Arabic consonants of interest to the present study are not available in this inventory, and thus, Yoruba speakers are expected to find it difficult to produce these consonants.

	Front	Central	Back
High	i		u
Mid	e ɛ		o ɔ
Low		a	

Table 4.24 The vowels of Yoruba (From Olmsted 1951)

	Voicing	labial	Labio-dental	Alveolar	Palatal	Velar	Labio-velar	Glottal
Stop	Voiced	b		d		g	gb	
	Voiceless		f	t		k	kp	
Affricate	Voiced			dʒ				
	Voiceless							
Fricative	Voiced							
	Voiceless			s	ʃ			h
Nasal	Voiced	m		n	ɲ			
liquid	Voiced			r l				
Approximant	Voiced	w			j			

Table 4.25 The consonants in Yoruba (From Olmsted 1951)

According to Olmsted (1951), the basic syllable shape of Yoruba is (C)V. Consonant clusters do not occur in the language nor do syllables or words that end with a consonant. Hence, it is predicted that Yoruba speakers will struggle in producing Arabic onset and coda clusters.

4.5 Comparison of the L1s and L2 Sound Systems

Comparisons between the L1 inventories of the FDHs in the present study and that of the target language is not done to highlight what substitutions the FDHs are predicted to have in their productions of Arabic segments. Rather, this was done to merely show the challenges the FDHs will face in perceiving or producing the consonants and consonant clusters of the target L2. With a group of foreigners with around nine L1 backgrounds, such as the one in the current study, it is hard to investigate every L1 and compare it to the L2 phonetically. Additionally, literature is scarce on some of the L1 backgrounds of the FDHs let alone a detailed description of the fine-grained phonetic details of specific sounds in their inventories. Based on this, a comparison between the sound inventory of Nizwa Arabic and that of the FDHs' L1s will be based on purely phonological description. Table 4.26 shows the consonant chart of Nizwa Arabic highlighting the consonants of interest to the current study that also turned to be absent in the L1 sound inventories of the FDH participants. Thus, we can confirm that FDHs' in the present study are comparable in terms of their L1s, at least with regard to the consonants of interest to the present study. Expressed differently, all consonants of interest to the present study are absent from all the L1 inventories of the FDH participants.

The target language and all the L1s of the FDH participants include the three cardinal vowels /i, u, a/ despite possible variation along the primary dimensions of the vowel space that pertains to language-specific acoustic configurations.

Table 4.27 shows whether the L1 phonology of the FDH participants involves onset and coda consonant clusters in comparison with the target language. Languages that permit consonant clusters are indicated with the symbol (√) while those that do not are indicated with the symbol (x). Languages are arranged in the table according to their syllable complexity (most complex to least). L1s that incorporate consonant clusters in loanwords are indicated with the (x) symbol as these are not part of the original L1 inventories of the learners and their production might still be affected by the L1 system. From this table, we learn that FDHs vary in whether their L1s incorporate onset consonant clusters. Some FDHs' L1s allow onset consonant clusters (e.g. Swahili, Luganda, Tagalog and Bengali) while others do not. However, FDHs' L1s are comparable in terms of the lack of coda consonant clusters. These differences conform to the universal principles of language. In Chapter 2, we learned that coda consonant clusters are more marked than onset consonant clusters due to their less frequent appearance in world languages.

	Voicing	labial	Labio-dental	Dental	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	Voiced	b			d		g			
	Voiceless				t		k	**q**		ʔ
Emphatic stop	voiceless				**tˤ**					
Fricative	Voiced			**ð**	z			**ʁ**	**ʕ**	
	Voiceless		f	**θ**	s	ʃ		**χ**	**ħ**	h
Emphatic fricative	Voiced			**ðˤ**						
	Voiceless				**sˤ**					
Nasal	Voiced	m			n					
Liquid	Voiced				ɾ l					
Approximant	Voiced	w				j				

Table 4.26 The sound inventory of Nizwa Arabic highlighting the consonants of interest that are absent from the L1 inventories of the FDHs'

Language	Initial CC	Final CC
Nizwa Arabic	√	√
Bengali	√	x
Swahili	√ (restricted)	x
Luganda	√ (restricted)	x
Tagalog	√ (restricted)	x
Yoruba	x	x
Oromo	x	x
Indonesian	x	x
Sinhala	x	x
Telugu	x	x

Table 4.27 A comparison of the syllable structure of the sound systems under investigation with regard to the existence or non-existence of onset and coda consonant clusters

4.6 Conclusion

This Chapter provided an overview of the consonants, vowels and consonant clusters of interest to the current study and discussed the rational for choosing them as well as reviewed literature on the articulatory, acoustic and phonological features that classify them. It discussed the acoustic correlates of fricatives and stops in order to establish the most prominent cues that NSs might use in case of hyperarticulation in FDH-directed speech. Furthermore, the Chapter presented the sound inventories of the target language (Nizwa Arabic) and those of the L1s of the FDH participants. Specifically, it shed light on the vowel and consonant phonemes that exist in each language as well as the syllable structure that is part of the phonology of these languages.

Now that the linguistic variables of interest have been illustrated and the sound inventories of L1s and L2 have been described, the following Chapter will present the methodology of the present study.

Chapter 5. Methodology

5.0 Introduction

This Chapter will start by restating the research questions of the study. This will be followed by a presentation of the hypotheses set to predict answers for the research questions based on the literature. A description of the participant population will then be provided. This will be followed by a description of the data collection tools used to obtain the data required for the study. A description of the data transcription, labelling, acoustic measurements, and statistical design will then follow before the chapter is concluded.

5.1 Research Questions and Hypotheses

As stated in Chapter 1, this study aims to answer the following questions:

Q1) What (if any) systematic adaptations are made in consonant, consonant cluster and vowel productions in FDH-directed speech? To what extent are these adaptations similar to or different from those reported in IDS, FDS, and clear speech research?

Q1a) Does FDH-directed speech contain hyperarticulation/enhancement of sound contrasts?

Q2b) Does LoR play a role in any adaptations made in speech to FDHs?

Q2) To what extent are FDHs perceptually sensitive to the phonemic contrasts of their L2 Arabic? What social and cognitive factors (independent variables) are responsible for any variation present in the listeners' performance?

Q3) How accurate (or target-like) are FDHs at producing Arabic consonants and complex syllable structures? What social and cognitive factors (independent variables) are responsible for any variation in the accuracy level of FDH's production skills?

I formulated the following hypotheses to answer the research questions above based on the literature presented in the previous Chapters. Hypotheses are arranged according to their relevance to each research question.

Q1)

H1. NSs will simplify complex consonants and produce sounds that are familiar to FDHs (e.g. /ʁ/→/g/; /q/→/k/).

H2. FDH-directed speech will involve less marked onset and coda consonant clusters.

The first two hypotheses are based on evidence from foreigner talk and baby talk discussed in Chapter 3 as well as based on differences between the L2 and L1s discussed in Chapter 4.

H3. NSs will hyperarticulate their speech acoustically and prosodically to enhance its clarity for their FDH interlocutors.

H3a. NSs will enhance the contrast between plain/emphatic consonants by hyperarticulating the prominent acoustic correlates that the NSs in the current study use to distinguish emphatic from plain consonants.

Based on literature of Arabic dialects (Al-Masri and Jongman 2003; Jongman *et al.* 2011; Al-Tamimi 2017) and literature on the acoustic correlates of stop and fricative consonants (Halle 1957; Fant 1960; Stevens and Blumstein 1978; Kewely-Port 1983; Jongman *et al* 2000; Kang and Guino 2008), the acoustic features that the study will focus on are spectral moments, F1, F2 and F3 of following vowels, VOT (for stop consonants), duration (for fricatives) and intensity. The exact acoustic correlates that will be enhanced will be determined after data analysis is carried out.

H3b. Consonant intensity will be higher in FDH-directed speech.

H3c. Fricative duration will be longer and VOT of stop consonants will be longer in FDH-directed speech.

H3d. Vowel space area will be expanded in FDH-directed speech indicating hyperarticulation.

H3e. Vowel intensity will be higher in FDH-directed speech.

H3f. Vowel duration will be longer in FDH-directed speech.

H3g. *f*0 will have no change in FDH-directed speech.

Hypotheses on vowel properties come from research on FDS, IDS and clear speech research (Chapter 3).

H6. Simplification and hyperarticulation in FDH-directed speech will vary based on FDHs' foreign accentedness score, LoR and religion.

This hypothesis is based on evidence on the influence of L2 learners foreign accent in triggering foreigner talk and from evidence on the role of infants' linguistic development or age in triggering IDS.

Q2)

H7. FDHs' perceptual sensitivity towards Arabic phonemic contrasts will not reach native-like performance.

H8. Foreigners with longer LoR, and who have more formal schooling and who are literate in their L2 will be more sensitive to Arabic phonemic contrasts.

Hypotheses under Q2 are based on literature on speech perception (e.g. Best and Strange 1992; Iverson and Khul 1996; Werker and Tees 1999) and factors affecting L2 speech learning discussed in Chapter 2.

Q3)

H9. FDHs will fall short of target-like accuracy in producing Arabic consonants.

H10. FDHs with longer LoR, and who have more formal schooling and who are literate in their L2 will be more accurate at producing consonants and consonant clusters.

H11. FDHs will fall short of target-like accuracy in producing Arabic consonant clusters.

H11.a FDHs will be more successful at producing onset consonant clusters than coda consonant clusters due to the markdness of the latter.

H12. FDHs with longer LoR, and who have more formal schooling and who are literate in their L2 will be more successful at producing Arabic consonant clusters.

Hypotheses under Q3 are based on literature from L2 adult speech production and factors affecting speech learning discussed in Chapter 2.

To answer the research questions and test the hypotheses above, three tasks were used. The aim of the first task was threefold: first, to investigate the phonological and acoustic properties of speech directed to FDHs and compare these properties to those in speech directed to a native adult (ADS); second, to find out whether these properties vary as a function of the FDHs' Arabic language experience indexed by LoR; third, to find out whether any of these properties reflect modifications similar to those reported in studies of clear speech, IDS and FDS. The other two tasks were used to examine FDHs' language outcomes by testing their perception and production abilities of Arabic sounds. The aim of the perception and production tasks were twofold: first, to examine the extent to which learners' sensitivity to Arabic phonemic contrasts and production of complex Arabic consonants and consonant clusters is native-like; second, to find out the role of the social and cognitive factors in their acquisition of these linguistic aspects.

The following sections will provide a description of the research design adopted in this study followed by a description of the sample.

5.2 The Research Design

This study adopts a mixed research design whereby quantitative and qualitative analyses were used to obtain findings and conclusions (DeMatteo 2005).

5.3 The Sample

Before providing a description of the sample, it is important to note that female FDHs were considered a vulnerable group when I applied for ethical approval for the study. Therefore, I provided several documents to justify that this group of participants would not be coerced to participate in this study. After ethical approval was obtained, and in order to decide on the sampling method, I distributed a questionnaire to 50 Omani families in the town of Nizwa who employ FDHs, briefly describing my study and asking them some background information about their FDHs (e.g. their L1s, religion, LoR) as well as whether they would be willing to participate in the study.[12] This was done to investigate whether there would be enough FDH participants to constitute a homogenous group with regard to their L1. Of those employers who were willing to participate, the L1 backgrounds of their FDHs were so diverse. Therefore, the L1 of the FDHs was not taken as a factor in the selection of the FDH participants. An additional aim in recruitment was to find FDH participants that varied in their LoRs in Oman or another Arabic-speaking country. However, it was hard to apply a strict criterion as to particular LoRs that could be divided later into two well-defined groups (short LoR and long LoR). This was also abandoned as a factor in selection. The method that was used to find participants was the "friend of a friend" technique proposed by Milroy (1980). I chose this method because as Schilling (2013: 192) said:

> We all feel more comfortable meeting people with whom we share an acquaintance or friend than we do meeting complete strangers, and often instant trust is conferred upon a new acquaintance when we know they are friends with a person we already know and trust.

Families do not trust strangers to make direct contact with their FDHs. Therefore, this method was the best to gain the trust of the employers and to then approach their FDHs for consent.

Twenty-two female native Omani-Arabic speakers and 22 female FDHs participated in this study. The former were the employers of the FDHs and served to provide the input data elicited in task 1 of this study. The employers were adults from the town of Nizwa in Oman, all spoke the Nizwa dialect of Omani Arabic and were of comparable educational levels (college students or graduates). The age of these participants was between 19 and 50 years (mean = 34). They will be referred to as "native speakers (NSs)" henceforth. To elicit ADS, I participated as the NS interlocutor. I was 32

---

[12] The Nizwa region has neighbouring suburbs that have their own dialectal features. Participants in the current study were recruited from within the borders of Nizwa city where reidents share similar dialectal features.

years old at the time of testing, originally from the town of Nizwa and spoke the Nizwa dialect of Omani Arabic. All participants reported having no hearing or articulatory problems.

The NS participants were asked to read a participant information sheet and sign a consent form prior to participation (Appendix 1). Similarly, FDHs were asked to do the same but the researcher read for them the participant information sheet and the consent form in a simple, clear language (Appendix 2).

Prior to the tasks, the FDH participants also provided background information, in response to questionnaire items which the researcher read to them and then recorded their answers (Appendix 3). In order to obtain a more precise account of the overlapping influences contributing to the learners' performance in the L2, the questionnaire included items that were related to five variables:

1. Age of arrival (AoA) in the Arabic-speaking world.
2. Length of residence: LoR in the Arabic-speaking world and in Oman
3. Native language education: years of formal schooling in the L1
4. L2 Literacy: ability in reading or writing in Arabic
5. Religion (i.e. Muslim vs. non-Muslim): exposure to spoken and written Arabic.

As the questionnaire was based on self-reports, it is understood that for L2 literacy, precise information might have not been obtained in comparison with, for example, studies that focus on literacy and measure it specifically. For native language education variable, it was not possible to measure the participants' literacy in these languages. However, they all reported being able to read and write in their L1s. Moreover, the fact that educational systems differ around the world made it impossible to provide the participants with a pre-defined set of educational stages to choose from (e.g. primary, secondary…etc.). Therefore, participants were asked to provide the total number of years they had spent at school. Not only was this easier for the current participants to express but was also easier to use as a continuous variable in quantitative analysis. Because native language literacy was not tested, educational level is used as a proxy for literacy in the present study.

It is also essential to note that the religion and the L2 literacy variables can overlap as a participant whose religion is Islam, for example, is more likely to be literate in Arabic via recitation of the Qur'an than a participant who is not a Muslim. Recitation of the Qur'an must be done correctly and fluently (Muhammed *et al.* 2012). Learning the Qur'an essentially involves learning the Arabic alphabet as well as learning *makhārij al-ḥurūf,* the properties of sound articulations (Supriyadi and

Julia 2019). Therefore, Qur'an tutors are expected to be experts in the principles of pronunciation and have knowledge of *ahkam al-tajweed*, rules for recitation (Damer *et al.* 2017). Recitation of the Qur'an may have started in different contexts including formal setting (State schools), Qur'an institutions or private tutoring, depending on the local policy of the FDHs' home country. Other than the recitation of the Qur'an, a Muslim can get exposed to Arabic during worship via prayers and educational lectures.

In addition to the variables above, participants were asked about the Arabic-speaking countries they had worked at prior to migrating to Oman due to the diversity of the Arabic dialects that might have affected the production patterns of the participants.

Table 5.1 presents the details of the background information of the participants and Table 5.2 presents the mean values of the numerical variables such as LoR. The FDHs came from nine L1 backgrounds (Swahili (5), Indonesian (2), Sinhala (4), Tagalog (5), Bengali (2), Telugu (1), Luganda (1), Yoruba (1), Oromo (1)). [13] They migrated to the Arabic-speaking world as adults (mean AoA= 27.27) and they had varying Arabic experiences based on their LoR that ranged from 0.7 to 21 years (mean LoR= 6.23). The mean of their LoR in Oman was 2.36 years. Some of them had worked in different Arabic-speaking countries before moving to Oman (e.g. Gulf countries and Lebanon). Some of them had never been exposed to Arabic before migration while others had had access to Arabic via Islam and recitation of the Qur'an. They all reported that they had been addressed mainly in Arabic by the family members of the household(s) they had lived in and worked for. Other than Arabic, some of them reported having some knowledge of other dialects and languages (mainly English) spoken in their countries of origin. It was however hard for the subjects to describe their proficiency in English language other than saying that they knew a few words and phrases that they used sometimes when they found it challenging to communicate in Arabic.

[13] The number between brackets indicates the total number of domestic helper participants with that L1.

Participant Code	Age at Testing (AoT)	Nationality	AoA to Arab World	LoR Oman (years)	LoR Arab World (years)	Arab Countries visited	L1(s)	L2(s)	Education	Literacy in Arabic	Religion
Lu	26	Uganda	26	0.7	0.7	None	Luganda	English	4	None	Muslim
Sw5	23	Tanzania	22	1	1	None	Swahili	English	6	None	Muslim
Yo	32	Nigeria	31	1	1	None	Yoruba	English	12	None	Non-Muslim
Sw1	19	Tanzania	16	0.16	1.75	Saudi	Swahili	English	10	Reading Qur'an & writing	Muslim
Ta1	23	Philippines	21	2	2	None	Tagalog	English	5	None	Muslim
Sw2	27	Tanzania	25	2	2	None	Swahili	English	7	None	Muslim
Ta2	22	Philippines	20	2	2	None	Tagalog	English	9	Reading Qur'an	Muslim
Sw3	21	Tanzania	19	2	2	None	Swahili	English	9	none	Muslim
Sw4	30	Tanzania	28	0.83	2.4	UAE	Swahili	English	7	None	Muslim
Si1	46	Sri Lanka	40	0.66	2.66	UAE, Kuwait	Sinhala	English	12	Reading Qur'an	Muslim
Be1	27	Bangladesh	24	3	3	None	Bengali	None	5	Reading Qur'an	Muslim
Te	38	India	34	1.8	3.8	Kuwait	Telugu	None	0	None	Non-Muslim
Or	27	Ethiopia	23	4	4.3	Saudi	Oromo	Amharic & English	7	Reading Qur'an	Muslim
Ta3	37	Philippines	27	0.7	5.7	Kuwait, UAE	Tagalog	English	12	None	Non-Muslim
Be2	30	Bangladesh	24	1.8	6.5	UAE	Bengali	None	0	None	Muslim
Ta4	30	Philippines	22	0.9	7.9	Saudi	Tagalog	English	12	Reading Qur'an	Muslim
Ta5	44	Philippines	35	0.3	8.3	Saudi	Tagalog	English	7	None	Muslim
In2	37	Indonesia	27	3	14	Saudi	Indonesian	English	4	Reading Qur'an	Muslim
In1	46	Indonesia	36	6	10	Kuwait, Saudi, UAE	Indonesian	None	3	Reading Qur'an	Muslim
Si2	46	Sri Lanka	33	4	13	Lebanon	Sinhala	None	10	None	Non-Muslim
Si3	48	Sri Lanka	28	10	20	Saudi, Kuwait	Sinhala	None	5	None	Non-Muslim
Si4	49	Sri Lanka	28	4	21	Kuwait, Saudi	Sinhala	None	4	none	Non-Muslim

Table 5.1 Background characteristics of FDHs from the questionnaire

Variable	Range (years)	Mean
Age of arrival	16-40	27.27 years
Age at testing	19-49	33
LoR Middle East	**0.7-21**	**6.23 years**
Formal schooling in the L1	**0-12**	**6.81years**

Table 5.2 Mean values of numerical background information of FDHs' (the highlighted information shows the variables examined in the present study)

The participants were also classified based on their religion as Muslims or non-Muslims. 16 of them reported being Muslims while six reported having a different religion than Islam. With regard to Arabic, all of the participants reported to be able to speak Arabic with varying degrees of proficiency depending on their experience. Eight of them reported to be able to recite the Qur'an, and only one reported to have some basic knowledge in writing along with recitation (Sw1). It should be noted that when classifying FDHs based on their Arabic literacy, only Muslim FDHs who reported being able to read in Arabic via recitation of the Qur'an were considered as literate. Those who reported not to be able to read in Arabic or recite the Qur'an were considered as non-literate in Arabic even if they were exposed to Arabic during other rituals of worship.

As to their exposure to Arabic after migration, the situation might differ considerably from one FDH to the other depending on the employer family. In most cases, exposure to Arabic can be highly limited to the household. In other words, communication between FDHs and members of the employer family mostly revolve around home chores and daily responsibilities the FDH is expected to accomplish. Additionally, input FDHs are exposed to mostly come from the female employer and her children rather than from the male employer or other male adults in the family. FDHs can also get exposed to foreign accented Arabic from other FDHs or NNSs around them, though this might depend on how often FDHs get in contact with such a group of NNSs.

The following section will describe the four tasks used to collect data for the current study. For each task, a discussion of the stimuli, task design and procedure will be provided.

5.5 Task 1: Spot the Difference to Examine FDH-directed Speech

In order to elicit data to investigate speech directed to FDHs, a spot the difference task was used adopting resources similar to those used in the diapixUK task (Baker and Hazan 2011). In this task, pairs of speakers engage in a spot the difference problem-solving task. Each speaker is given a different version of the same picture, and the two speakers have to interact with each other collaboratively to find differences (in this case 12) between the two pictures (Van Engen *et al.* 2010; Baker and Hazan 2011). This kind of task allows the speaker to engage in a constrained yet natural interaction with the interlocutor, which helps to elicit data that is specific but less contrived compared to other tasks that require reading or entail specific texts (Knoll and Scharrer 2007). Although natural interactive speech recordings are challenging to process and analyze, there is a higher possibility to generalize results from these data to natural speech, which makes this challenge worth solving (Baker and Hazan 2011). This is because the level of difficulty of the task that is represented by the rich content of the pictures allows for complex speech structures to be generated from the participants.

This task was useful for the current study because it can be used to examine different aspects of speech used during interaction such as acoustic-phonetic properties, speech rate and vowel space (Baker and Hazan 2011). Moreover, it allows the examination of specific segmental analyses via manipulating the picture differences to be based on keywords that contain the target segments. Thus, the benefit of specifying speech targets is to reduce potential acoustic variability that could result from changes in the phonetic context across different speech styles (Ferguson and Kewley-Port 2002; Krause and Braida 2004). The task could be manipulated so that one participant talks more than the other which can be done by instructing the former to 'take the lead' in the interaction (Baker and Hazan 2011: 763).

5.4.1 Stimuli

Prior to designing the task for this study, three lists of keywords that contain the target segments of interest were created. The first list consisted of 24 Arabic content words that contained the target consonants (Table 5.3). To ensure the FDHs successfully engaged in the task, the keywords were chosen to represent familiar objects in the household. Each target consonant appeared word initially in two keywords. The keywords were not comparable in terms of word stress as it was hard to find words that all shared the same stress pattern and at the same time were relevant to the picture themes as well as not cause a barrier to communication. The second list contained 10 content words with initial and final two-member consonant clusters (Table 5.4). The clusters were chosen randomly but had to be in a content word relevant to the household. The third list consisted of 9 content words that contained the three Arabic corner vowels illustrated in table 5.5 (4 are repetitions of keywords from tables 5.3 and 5.4). All vowels appeared in stressed syllables. This was to control for the effect of stress. To this end, a total of 43 stimuli were used in this task.

Target Consonant	Keywords	Target Consonant	Keywords
/χ/	/χass/ ‘lettuce’ /χaruːf/ ‘sheep’	/θ/	/θoːm/ ‘garlic’ /θalaːθeh/ ‘three’
/ħ/	/ ħabɪl/ ‘rope’ /ħaliːb/ ‘milk’	/ð/	/ðurah/ ‘corn’ /ðahab/ ‘gold’
/q/	/qalam/ ‘pen’ /qɪtˤaːr/ ‘train’	/ʁ/	/ʁarʃeh/ ‘bottle’ /ʁassaːleh/ ‘washing machine’
/ðˤ/	/ðˤarf/ ‘envelope’ /ðˤbaːb/ ‘flies’	/ʕ/	/ʕalam/ ‘flag’ /ʕankbuːt/ ‘spider’
/sˤ/	/sˤaħan/ ‘plate’ /sˤaabuːn/ ‘washing liquid’	/s/	/samak/ ‘fish’ /sajjaːrah/ ‘car’
/tˤ/	/tˤaːwleh/ ‘table’ /tˤamaːtˤ/ ‘tomatoes’	/t/	/taːg/ ‘crown’ /talafuːn/ ‘telephone’

Table 5.3 Target Consonants and Keywords

Onset consonant cluster	Keyword	Coda consonant cluster	Keyword
/st-/	/staːrah/ ‘curtain’	/-nt/	/bɪnt/ ‘girl’
/dg-/	/dgaːg/ ‘chicken’	/-bz/	/χubz/ ‘bread’
/lħ-/	/ lħam/ ‘meat’	/-lb/	/kalb/ ‘dog’
/gd-/	/gdaːr/ ‘wall’	/-rf/	/ðˤarf/ ‘envelope’
/kt-/	/ktaːb/ ‘book’		
/ðˤb-/	/ðˤbaːb/ ‘flies’		

Table 5.4 Target consonant clusters and keywords

Target vowel	Keyword		
/iː/	/fiːl/ ‘elephant’	/tiːn/ ‘fig’	/ħaliːb/ ‘milk’
/aː/	/baːb/ ‘door’	/taːg/ ‘crown’	/ktaːb/ ‘book’
/uː/	/fuːl/ ‘chickpeas’	/tuːt/ ‘berries’	/χaruːf / ‘sheep’

Table 5.5 Target vowels and keywords

5.4.2 Task design

Since the FDHs in this study had been exposed to a language that is highly limited to the household, it was necessary to design the task in a way that would allow the NS and FDHs to engage in a familiar, typical conversation. Therefore, the pictures for the task were designed particularly to reflect some household scenes that were assumed to be familiar to the FDHs and easy for the NS to describe. The task was to some extent designed in a way that took into consideration the criteria used in designing the DiapixUK task in terms of style, complexity and number, and type of differences (Baker and Hazan 2011).

The task consisted of six picture pairs with three different scenes: kitchen, living room, and bedroom, with two different pairs per scene (Appendix 4). The scenes were graphically designed in a cartoon style by a graphic designer. They contained objects that represented specific keywords (the keywords in tables 5.3, 5.4 and 5.5) and some distracters. The differences between each picture pair depicted a change in the color, number and presence or absence of an object across the two pictures (e.g. a red apple in X vs. a green apple in Y or three cushions in X vs. two in y). When the difference was based on the presence or absence of an object, the picture that the NS had was always the one that had the object present while the picture that the FDH had was always the one that had the object absent, especially if the object represented a keyword. This is because it was essential to ensure that the pictures the NS had contained all the stimuli of interest to the study, and thus it was not possible to remove an object that represented a target word from the NS's picture. In prosodic terms, this could also help to maintain sentential stress on the target words. Each pair contained twelve differences that were related to some of the keywords. This was the number of differences used in previous studies and is believed to provide sufficient time for the subjects to generate more tokens and not too long for them to be bored (Baker and Hazan 2011).

The rationale behind creating two different pairs per scene was to use one pair to elicit FDH-directed speech and the other pair to elicit ADS, and thus to avoid repetitiveness. The two versions of pictures per scene had the same keywords from tables 5.3, 5.4 and 5.5 but differed in the type of differences (e.g. color, number, etc.), the order of objects and the kind of distracters. This was to ensure the NS produced the same keywords in her description of the pictures to both the FDH interlocutor and the adult NS interlocutor, which could have yielded comparable speech samples across participant groups while reserving the complexity of the task. The keywords were distributed across the three scenes based on how one keyword connected to a particular scene more than the other scenes. For instance, keywords that represent vegetables (e.g. /χɪss/ 'lettuce', /tˤamaːtˤ/ 'tomatos'/) associate more with a kitchen scene compared to other scenes; whereas keywords that represent stationery (e.g. /ktaːb/ 'book', 'qalam' 'pen') associate more to a living room or a bedroom scene than a kitchen scene.

5.4.3 Procedure

The NS and the FDH were accompanied by the researcher to a quiet room in the house and each given three pictures of three different scenes arranged in the same order. They were instructed to sit opposite each other and try to find 12 differences between each picture pair without seeing each other's pictures. The NS was encouraged to take the lead and negotiate with her FDH the differences by asking her different questions and making descriptions of the scenes. She wore a Sennheiser radio microphone and the interaction between her and the FDH was recorded using Edirol digital recorder R-09HR by Roland with a sampling rate of 44.100 Hz. The recording mode was WAV-16bit.

The NS was then invited to participate in a similar task using the other set of picture pairs with the researcher. It is important to point out that as a control measure, 11 of the NSs were asked to play the game with their FDHs first and then with the researcher while 11 of them were asked to play the game with the researcher first and then with their FDHs. This yielded two condition orders: (1) order 1 (FDH-directed speech/ADS), (2) order 2 (ADS/FDH-directed speech). This was done to control for any condition order effect that could influence the way the NS would interact in the task.

5.5 Task 2: AX Discrimination Task for FDH Perception

Two tasks are often used to show categorical perception in adults: identification and discrimination Kuhl 2007). In the current study, an AX forced-choice discrimination paradigm was used to elicit FDHs' sensitivity to different Arabic consonantal contrasts. In this kind of task, participants are presented with two stimuli in sequential order, the second stimulus is either the same as the first (AA) or different (AB) (Strange and Shafer 2008). Participants need to compare the two stimuli and judge whether they are the 'same' or 'different'. It should be noted that the ABX paradigm, in which the last item in a triad is judged to be the same as the first or second items is more often used in discrimination tasks (Brain 1980). However, that paradigm appears to have the disadvantage of requiring overt categorization responses and heavy memory load (Pollack and Pisoni 1971). Since the present study was interested in examining FDHs' auditory sensory discrimination skills only, the AX task was appropriate because it constitutes the lowest memory load and stimulus uncertainty compared to other discrimination tasks (Strange and Shafer 2008). Thus, participants will focus their attention on judging the two contrasts rather than remembering what items they had heard. Furthermore, the AX task was easier to describe to the FDH participants given their limited Arabic capacity. The AX task was also used instead of a categorization task due to the fact that most of the FDHs were non-literate or low-literate in Arabic and it would have been hard for them to deal with a more complicated task that might require a computer version, for instance.

One disadvantage of the AX discrimination paradigm is that it has a high probability of response bias. To clarify, participants in such a task are likely to respond "same" or "different" randomly when they are unsure of the answer. Thus, counting only % correct on the different pairs is not an accurate measure of subjects' responses. To account for the FDHs' accuracy, the current study used sensitivity measures from Signal Detection theory (Macmillan and Creelman 1991).

5.5.1 Stimuli and task design

A list of 16 Arabic consonant contrasts was created. The contrasts, which were produced by me, contained all the target consonants included in task 1 (see the end of this section on how recording of the test items was carried out). The phonemic pairings were created based on their potential confusability for the listeners which was determined from the following criteria:

1. Pairs contrasting in voicing, manner or place features; e.g. /θ/-/ð/, /tˤ/-/t/, /χ/-/ħ/ respectively.
2. One stimulus in a pair is a likely variant of the other in FDHs' productions. For example, /q/-/k/ might be perceptually confusable because /k/ is likely to replace /q/

in FDHs' productions. This was based on my own observations of FDHs' substitution patterns.[14]

In addition, two more contrasts were used as control /r/-/l /, /r/-/w/. Given that the FDHs' L1 inventories all included a /r/, a /l / and a /w/, it was likely that all FDH participants would detect these contrasts as different (Aoyama *et al*. 2004). With the control pairs, a total of 18 consonant contrasts was included in the test.

Using an AX discrimination task, four test items were created for each of the 18 contrasts (AA, AB, BA, BB) yielding a total of 72 test trials. Each test trial consisted of two monosyllabic pseudo words containing the contrasting sounds in the context Caːn (where C is a consonant), for instance, /saːn- sˤaːn/. 36 trials contained consonants that were acoustically different while 20 trials contained consonants that were acoustically identical (see appendix 5 for all trials before randomization). The use of pseudo words instead of real words minimizes the influence of the lexical representations of real words that could affect the listeners' discrimination ability (Strange and Shafer 2008). Moreover, using real words may yield inaccurate interpretations of performance when comparing inexperienced and experienced learners (ibid.). Expressed differently, less experienced FDHs might be unfamiliar with some of the words, which would give an advantage to the more experienced FDHs. The phonetic context in which the target consonant was embedded was chosen to contain sounds (in this case "C**an**") that exist in most world languages (Maddieson and Disner1984). This was done to reduce any effect of native language phonotactics on listeners' discriminative abilities. Table 5.6 shows all consonant contrasts used in this task and the phonetic environment they appeared in. It should be noted that when all the sounds of interest were inserted in this phonetic context, it appeared that for some words what was potentially meant to be a pseudo word turned out to be a real Arabic word. For example, the sounds /χ/ and /sˤ/ made the words /χaːn/ and /sˤaːn/. These are real Arabic words meaning 'he betrayed' and 'he protected' respectively. It was, however, deemed unlikely that the FDH subjects in the present study were familiar with such words as they are very likely to be infrequent in the input context they are exposed to.[15] In addition, as we can see from Table 5.6, most of these words come from MSA and being a speaker of Nizwa Arabic, I find such

[14] My observations were not based on all L1 languages of the FDHs but were general observations made from my experience with female FDHs.

[15] Frequency is established based on the researchers' own observation as a member of the speech community that FDHs work for and as a speaker of the dialect of this community.

words infrequent if not absent in the dialect except for ‘kaːn’ ‘he/it was’ which is used regularly in everyday communication.

Consonant contrast	Stimuli
/Ɵ/-/ð/	/Ɵaːn/ - /ðaːn/
/χ/-/ʁ/	**/χaːn/** - /ʁaːn/
/ħ/-/ʕ/	/ħaːn/ - /ʕaːn/
/tˤ /-/t/	/tˤaːn/ - /taːn/
/ðˤ/-/ð/	/ðˤaːn/ - /ðaːn/
/s/-/sˤ/	/saːn/ - **/sˤaːn/**
/q/-/χ /	/qaːn/ - /χaːn/
/ħ /-/h/	**/ħaːn/ - /haːn/**
/k/-/q/	**/kaːn/** - /qaːn/
/χ/-/ħ /	/χaːn/ - /ħan/
/r/-/l/	**/raːn/ - /laːn/**
/r/-/w/	/raːn/ - /waːn/
/χ/-/k/	/χaːn**/-/kaːn/**
/Ɵ/-/s/	/Ɵaːn/-/saːn/
/Ɵ/-/t/	/Ɵaːn/-/taːn/
/ð/-/z/	/ðaːn**/-/zaːn/**
/ð/-/d/	/ðaːn/-/daːn/
/ʁ/-/g/	/ʁaːn/-/gaːn/

Table 5.6 Arabic consonant contrasts and stimuli used in the perception experiment (the highlighted stimuli are real Arabic words)

16 trials were excluded from the list as they were repetitions of existing trials. This reduced the number of trials to 56. To give an example, /θ/ was paired twice, once with /s/ and another with /t/ because /t/ and /s/ are likely variants of /θ/ in NNS' productions as we have seen in Chapter 4. When the four test trials were created for each of these pairs, one test item was repeated for both contrasts. Table 5.7 shows that trial 8 (/θan/-/θan/) is repetitive as it is a resemblance of trial 4. Hence, to avoid repetition, one of these test trials was excluded.

Contrast (AX)	Test trials			
	1	2	3	4
/θ/-/s/	/θaːn/ /saːn/	/saːn/ /θaːn/	/saːn/ /saːn/	/θaːn/ /θaːn/
	5	6	7	8
/θ/-/t/	/θaːn/ /taːn/	/taːn/ /θaːn/	/taːn/ /taːn/	/θaːn/ /θaːn/

Table 5.7 Example of trials that were deleted

When all trials were created, they were submitted to an online randomization software (RANDOM.ORG) to ensure that the four test items for each contrasting pair were not following each other (see appendix 6 for all trials after randomization).

I produced the test items which were transcribed using the IPA system to facilitate pronunciation.[16] The stimuli were recorded in a sound treated lab using high quality recording equipment and a microphone. Another native Omani speaker trained in linguistics listened to the recorded stimuli for a reliability check. She was a 35-year-old IPhD student in linguistics at Newcastle University and spoke the same dialect as the NSs in this study. She confirmed that all recorded instances were clearly articulated and time intervals between each test item and the other were equally the same for all test items and as specified in the present study.

[16] IPA stands for the International Phonetic Alphabet. It is the most widely used system for alphabetic transcribtion (Rajab 2013).

5.5.2 Procedure

After performing task 2 and in the same quiet room, each FDH was presented with the aural stimuli over headphones at a comfortable volume level using a MacBook laptop. They were instructed, in Arabic, that they had to decide if the two test items they were about to hear were the same or different. They had to give their responses to the researcher who manually entered them on an answer sheet designed specifically for this task (Appendix 7). It was not possible to use a computer version for responses to this task due to potential difficulty the participants might have faced dealing with such programmes given their limited Arabic capacities and for some limited formal education. An inter-stimulus interval (ISI) of 1 second was used between each word in a comparison pair. Longer ISI is shown to facilitate phonemic discrimination rather than phonetic discrimination of non-native contrasts (Werker and Logan 1985). The participants were allowed to replay a trial as often as they wished, but they could not change an answer already given (Guion *et al.* 2000). All 56 trials were presented in two blocks during a twenty-minute session. A three- minute interval separated the two blocks.

To ensure that the participants understood the task procedure, they were presented with a familiarization task prior to the experiment. They were trained to listen to and judge two contrasts (/ʃ/ vs. /s/, /t/ vs. /d/) and were given immediate feedback about the correctness of their responses. The contrasts presented in the familiarization task were different from those used in the real test to avoid providing the participants with help on the target stimuli (Beddor and Gottfried 1995). Adopting a similar approach to Aoyama *et al.* (2004), the participants had to respond correctly to at least 90% of the stimuli in order to proceed to the actual task. If a participant did not reach this standard, the practice task was repeated up to 4 times or until they met the standard. Two of the participants (Sw2 and Te) were excluded from taking the real task as they failed to meet this standard on the familiarization task. Thus, only 20 FDH participants carried out this task.

10 native Omani speaker females were recruited as a control group in this task. They all had a comparable educational background and were between 19 and 40 years old. A similar procedure was used to elicit NS responses to the same task. However, the NSs were given the answer sheet to record their own responses and were not presented with the familiarization task prior to the experiment. In addition, NSs were not allowed to replay a trial but were rather encouraged to answer all of the test trials in one go.

5.6 Task 3: Picture-Naming for Arabic Production

The purpose of the picture-naming task was to elicit speech samples from FDHs in order to examine the characteristics of their production of Arabic consonants and syllable structure. To recapitulate, the present study was interested in uncovering the extent to which the target consonants and consonant clusters were produced accurately by the FDHs and to examine the factors that pertain to any differences in the accuracy level of the speakers. A picture-naming task is appropriate and ecologically valid for this purpose because it elicits data that is typical of natural speech (Trofimovich and Baker 2006). In addition, the task could be controlled to involve all sounds of interest to the study as well as control their phonetic context (Edwards and Beckman 2008).

5.6.1 Stimuli

The keywords in Tables 5.3 and 5.4 which contain the target consonants and consonant clusters were used to elicit single words from the particpants. The study considered keywords that contained /t/ and /s/ as controls but also added two more keywords beginning with the sound /k/, namely */kabat/* ‘closet’ and */kabriːt/* ‘match’. Accordingly, the total target words used in this task were 36.

5.6.2 Task Design

36 pictures that represented the stimuli keywords were used (Appendix 8). The pictures illustrated home objects and hence were assumed to be familiar to most FDHs.

5.6.3 Procedure

After taking the discrimination task, I asked each FDH to name the objects in the pictures presented to them in a MacBook laptop. The same high-quality recorder and microphone used in task 1 were used again here to record the participants’ productions. When a participant struggled to name an object, a delayed repetition technique was used (Ratner 2000; Guion *et al.* 2000). I, first, produced a short question (prompt) and then gave an answer to the prompt. I, then, uttered the prompt again and the FDH participant had to repeat the response they had heard. For example, to elicit the stimulus word ‘/χaruːf/ ‘sheep’, the prompt-response sequence below was used.

(5.1)

(a) Prompt: what is this? (short pause)
(b) Response: this is /χaruːf/?
(c) Prompt: what is this?
(d) Response (FDH): …

The delay between the prompt and the participant's repetition (approx. 4 s) had the benefit of minimizing the effect of direct mimicry that could be involved in repetition tasks.

The following sections will provide a description of the data analysis process. This will include a description of data transcription and labelling. It will also cover the acoustic analysis process carried out on data collected from task 1. In addition, it will include the coding process and statistical tools used to obtain the results.

5.7 Data Transcription

Data from tasks 1 and 3 were transcribed manually using the speech analysis software Praat (Boersma and Weenink 2009). I used orthographic transcription in all contexts that involved the target words of interest. I used the ASCII Phonetic Alphabet, an alternative transcription system which, unlike the IPA, does not require the use of special fonts and allows typing the sounds on a computer easily. For example, the IPA symbols of the Arabic sounds ʕ and ħ are represented as *Qh* and *H* respectively in the ASCII phonetic characters. Labelling procedures for task 1 are described in 5.8.

I transcribed data from task 3 (picture-naming task) in Praat for ease of extraction of the target segments. I created several tiers, for example, the target sound, the IPA pronunciation, the actual pronunciation, whether production was correct or not and other comments. I, then, designed a Praat script to transfer all tier information into Excel files.

5.8 Labelling

After I transcribed all data of interest from task 1 orthographically, I moved on to do labelling for the segments of interest. I labelled all instances of the sounds of interest that were free of background noise, tongue stuttering and speaker overlap. Praat features allow for the creation of multiple tiers that the researcher could use not only to label the segments of interest but also to add any information essential for the description of these segments. Thus, the tiers allow providing information (e.g. segment environment) that will eventually accompany the results sent to data files and can largely assist subsequent statistical analyses (Owren 2008). The kind of information the researcher decides to add in the tiers depends on the features they intend to investigate. Therefore, it is essential to think ahead of the exact analyses one needs to carry out on the segments of interest. For example, for transcribed data from task 1, I created multiple interval tiers that defined the target words, their actual pronunciation by the speaker, their position in the phrase (at a prosodic boundary or not), the target sound of interest to the study, its actual pronunciation by the speaker and whether it was stressed or not. I also added a tier named 'Epenthesis', where I recorded whether a syllable structure contained an epenthetic vowel or not by writing yes or no. Also, I added a tier for stop consonants to define their VOT. Finally, I added an extra tier named ‘sound type’ to differentiate between the three-point vowels /aː, iː, uː/ and the vowels following the consonants of interest. Figure 5.1 provides an example of the tiers I used for labelling data from task 1.

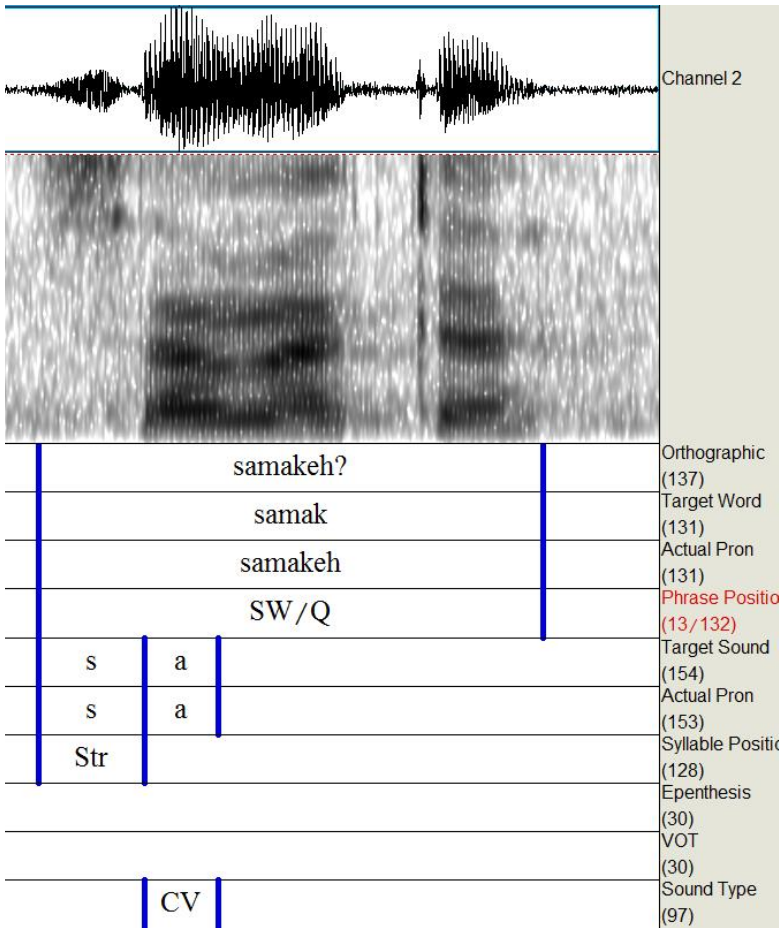


Figure 5.1 Tiers used for data labelling in Praat

Praat does not provide constraints on labelling. Therefore, the user is required to follow certain conventions in the labelling process (Owren 2008), and to be consistent with whatever convention one decides to adopt. For vowel labelling, I followed Al-Tamimi’s (2017) labelling protocol with some modifications. I determined the onset and offset of the vowels /a, aː, i, iː, u, uː/ by the first and last periodic pulse in the waveform and the vertical consistency of the first three formants in the 0-5500 Hz scale of the spectrogram. In cases where vowels were followed by a sonorant (e.g. in /fiːl/

‘elephant’, /tiːn/ ‘fig’), I determined the boundary position of the vowel based on intensity drop and visual assessment. Figure 5.2 illustartes an example of vowel labelling for the word /tiːn/ used in this study. [17]

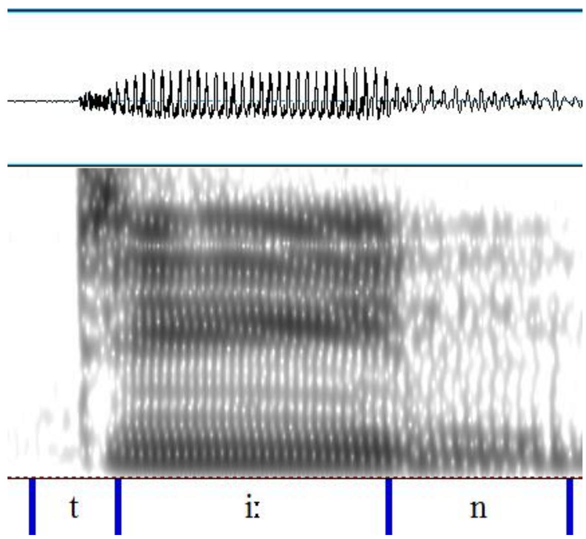


Figure 5.2 Example of target vowel labelling in Praat

I determined the boundaries for the plain/emphatic fricatives based on the first visible and audible frication (noise energy), including any silent interval which sometimes preceded the onset of the following vowel (following Khattab and Al-Tamimi 2008). For plain/emphatic stops, the onset of the sound was determined based on context. If the stop was preceded by a final sound of a previous word, I determined the boundary based on the first silent interval that immediately followed the preceding sound.[18] If, however, the stop was preceded by a long pause, the start of closure was impossible to determine and, therefore, I determined the beginning of the stop based on the first visible and audible frication. Hence, the measurement of such stops will only include the burst but not closure (Al-Ani 1970). I determined the onset of the burst based on the visible and audible frication but more precisely based on the first pointy peak in the waveform at 0.025-0.018 ms visibility. The offset of the stop was determined based on the onset of voicing which was established

[17] The figure shows a modified version of the original Praat’s textgrid (e.g. Figure 2). Only the tier that has the segment of interest is shown here and all other tiers have been removed.
[18] This is because the current study examined voiceless stops only.

based on the first visible pulse in the waveform and the vertical consistency of the first three formants of the following vowel.

5.9 Acoustic Measurements

I designed several Praat scripts to automatically perform acoustic measurements for all sounds of interest obtained from task 1. To examine the acoustic properties of vowels and whether FDH-directed speech is hyperarticulated, this study focuses on the following acoustic indicators: vowel space size, absolute formant frequencies (F1 and F2), fundamental frequency (*f*0), vowel duration and intensity, although it should be noted that the effect of speech style (FDH-directed vs. ADS) could be reflected along other acoustic dimensions (Bradlow and Alexander 2007). Measurements of all metrics (except duration) were obtained at the midpoint position of each vowel as this is the point that is the most stable and the least vulnerable to influences from neighboring segments. The configuration for formant calculation was obtained from a 25-ms Kaiser2 window with a 5-ms formant time step. Using the default Burg algorithm, a maximum of five formants was used for formant estimation with a maximum frequency of 5500 Hz for female speakers. A manual verification was then carried out on formant frequencies to check for any errors that might result from automatic extraction. Errors detected were corrected by hand. Once the formant frequencies were extracted, they were converted to the psychoacoustic Bark scale using the formula in 5.1 from Traunmüller (1990). Furthermore, sound duration was measured in millisecond and intensity in decibel.

(5.1)

Z_n= {26.81/ (1+1960/f_n0}-0.53, where Z_n is the Z value of a formant n and the f_n is a formant's frequency in Hz

To examine vowels following plain/emphatic consonants, measurements of formant frequencies (F1, F2 & F3) were obtained at the onset position of each vowel. A similar configuration to the one above was used and formants were transformed to the psychoacoustic Bark scale too. To examine the acoustic properties of the fricative plain/emphatic consonants, measurements of the first four moments (Centre of Gravity, Standard Deviation, Skewness and Kurtosis) were obtained from the mid position of the consonants. The configuration for moment calculation was obtained from a 20-ms Kaiser2 window following Jongman *et al.* (2007). For plain/emphatic stops, moments were obtained using a 15-ms kaiser2 window centered at the onset of the burst. This window size was chosen to ensure the least influence from neighboring segments.

5.10 Reliability

For reliability check of the segmentation process for data from task 1, I repeated 10% of the data (this was calculated for the overall target words segmented from this task). Results of the auditory analysis of all consonants showed 100% agreement judgment. I also repeated the labelling criteria for all target consonants that required acoustic measurements. Around 4 instances showed a significant deviation from initial labelling. These were then reviewed and corrected following the conventions used in the present study. For the picture naming task, 10% of the data was presented to a male native-speaker of Omani Arabic (speaking the Nizwa dialect) and asked to rate the consonants he heard as *target or non-target* as well as to transcribe the target words. Rating yielded 100% agreement judgment with the initial analysis with regard to how target-like the productions were. Transcription of the target consonants had 9 differences from the initial transcription especially with the kind of substitutions used by the speakers. Where differences were significant, the recording was reviewed again in the presence of both transcribers and an agreement over which segment was used was settled upon. However, this was not of importance to the current study as this study was not interested in exploring the kind of substitutions used by the speakers but rather the accuracy of their productions.

5.11 Statistical Design

All statistical analyses were carried out in R using several packages (R Core Team 2012). In the following sections, a description of the statistical models used, and their structure will be provided for every linguistic variable or task.

5.11.1 Consonants and consonant clusters in FDH-directed speech

I designed a Praat script to extract all instances of consonants and syllable structures of interest. After being transferred to excel files, tokens and percentage native-like production of the consonants of interest were generated in R. No statistical test was used for this part and alternatively results were discussed based on descriptive statistics as well as qualitative analysis based on spectrograms. For consonant clusters , to examine the effect of speech type (FDH-directed speech/ADS) on realization of these in onsets and codas, each cluster was assigned a value of 1 or 0 depending on whether there was deletion or not (for onsets) and whether there was epenthesis or not (for codas).[19] A generalized linear mixed effect model (GLMM) was then used to generate coefficients using the lme4 package (Bates *et al.* 2012). Speech type was used as a fixed effect. For random effects, I used speaker and

[19] I refered to modification of onset consonant clusters as ‘deletion’ and not ‘epenthesis’ because the phonological rule of syncope in Nizwa Arabic suggests that the short vowel is deleted in onset unstressed syllables.

target word as random intercepts. To examine the effect of LoR, LoR was treated as a continuous variable because it would have been misleading to divide FDHs into two groups based on their LoR without a strong rationale for this division. As we saw in Table 5.2 in Section 5.3, there was no clear cut off point that could provide two neat categories of LoRs. On a subset of the data only, that of FDH-directed speech, LoR as a fixed effect was centered before it was submitted to GLMM analysis.[20] For random effects, the same structure as the one above was used. To test whether any significant changes in vowel deletion or epenthesis were due to interlocutor order in carrying out the task (Table 5.8), the difference in consonant cluster realization between FDH-directed speech and ADS was estimated for each condition order separately using GLMM. Speech type was used as a fixed effect. The random effect structure was similar to the one above.

Condition Order 1	FDH-directed speech	ADS
Condition Order 2	ADS	FDSH-directed speech

Table 5.8 Condition order

I chose the generalized linear mixed effect analysis with random intercepts for three main reasons. First, Barr *et al.* (2013) recommend that when selecting a model for analysis, researchers should "keep it maximal" in terms of the random effect structure. Second, a model with random intercepts but not random slopes as random effects appeared to be the model that is more explanatory of the data after the Akaike Information Criterion (AIC) was applied for model selection; thus, this was the maximal random effect structure that fits my data (Symonds and Moussalli 2011). Third, my design included the need to register different measurements per speaker who produced multiple and sometimes repetitive target words that contained the target vowels and consonants. This violates the independence assumption that is a very essential requirement of the linear model (Winter 2013). To explain, multiple and repetitive productions from the same speaker cannot be regarded as independent from each other. We would expect each speaker to produce the target vowels with varying F1 and F2 for example. In consequence, this is going to be an idiosyncratic factor that affects all productions from the same speaker. Therefore, the model I chose for my analysis is one that accounts for these interdependencies by including random effects from speaker and target word.

[20] Prior to the application of many regression models, data centring is used as a pre-treatment analysis for continuous variables. Centring means "subtracting the sample mean from all input variable values" (Schielzeth 2010:104). This is essentially done to improve the interpretability of model coefficients (ibid).

5.11.2 Vowels in FDH-directed speech

I used the PhonR package (McCloy 2016) to plot the first and second formant frequencies (F1, F2) of the three Arabic vowels /iː, aː, uː/ as produced by NSs in speech directed to FDHs and in speech directed to the NS adult as well as to calculate the vowel space area of the two speech registers. For vowel plotting, I used the mean F1 and F2 values at midpoint of the vowel. To generate the vowel space area, I calculated vowel space size for each speaker separately using convex hull area using PhonR. I then used a linear mixed effect model (LMEM) using lme4 package (Bates *et al.* 2015) to examine the effect of the speech type on vowel space (hyperarticulation). Since the authors of lme4 package have not included p-values in the results of the LMEM, I used lmerTest package to obtain p-values (Kuznetsova *et al.* 2017). Vowel space was used as the dependent variable and speech type was used as the independent variable. For random effects, I used speaker as random intercept. A model with target word as random effect refused to converge. In addition, log likelihood tests selected a model with speaker as intercept as the optimal model. To examine the effect of foreign accentedness, LoR and religion on vowel space, an LMEM was used. The same random effect structure as the one above was used.

To examine changes in F1 and F2 in relation to speech style, LMEM was used for each vowel separately. F1 of all three vowels was examined while F2 of the front vowel /iː/ and the back vowel /uː/ were only examined following previous research (Ferguson and Kewley-Port 2002). The same model was used to examine the effect of speech style on intensity, duration and *f*0 of all three vowels. In the model, speech type was used as a fixed effect. For random effects, I used speaker and target word as intercepts.

To examine whether any changes in F1, F2, duration, intensity, *f*0 are the by-product of phrase boundary in which the target words appeared, each vowel was assigned a value of either 0 or 1 depending on where in a phrase boundary it appeared (at a phrase boundary or not) following (Miyazawa *et al.* 2017). Two linear mixed effects models were used to measure the size of effect, one including speech type and phrase boundary as fixed effects (full model) and another without speech type (reduced model). The full model and the reduced model were estimated for F1 and F2 values of each vowel separately as well as for duration, intensity and *f*0 metrics of all three vowels (16 models). The same random effect structure used above was used in these models. A Likelihood Ratio Test was used to compare the likelihood of the two models with each other and obtain the significance of one model over the other (Barr *et al.* 2013). To test whether there is any effect of condition order on any significant changes in F1, F2, duration, intensity and *f*0, LMEM was used in a similar way and using a similar fixed and random effect structure as the above model.

5.11.3 plain/emphatic consonants and the vowels following them

To examine the effect of emphasis and speech style on spectral moments, duration and intensity of plain/emphatic consonants in ADS and FDH-directed speech, LMEM was used. Sound type(plain/emphatic) and speech style were used as fixed effects. Also, interaction between sound type and speech style was included. Speaker was used as random intercept for random effects. Target word was not included in the random effect structure because a model with speaker only as random intercept was found to be more meaningful than one with speaker and target word according to AIC.In addition, a model did not converge when target word as random intercept was included in the model. As a follow-up analysis, *post hoc* pairwise comparisons were made using the *emmeans* package in R (Lenth and Lenth 2018). This package helps obtain "estimated marginal means (EMMs)" for a number of models as well as compute contrasts of EMMs and comparisons of slopes (Lenth and Lenth 2018: 3). Since the fitted LMEM model that I used contained interaction, I did the comparison of the speech styles separately for each sound to avoid making conclusions based on marginal means that are involved in interaction. The degrees -of-freedom method used was Kenward-roger with confidence level= 0.95. Only for significant modifications in any of the dependent metrics that the effect of condition order, phrase boundary and word stress were examined. This was done using LMEM with a similar random effect structure as above. On a subset of the data, that is of FDHs, LMEMs were used to examine the effect of foreign accentedness, LoR and religion on all metrics. L1 of the FDHs was not included in the model because a model with L1

5.11.4 Discrimination of L2 consonant contrasts

Data from the AX discrimination task was coded and tabulated manually in Excel. All background characteristics of the subjects were added, too. As mentioned in Section 5.5, counting % correct of the responses is not a meaningful measure. Thus, the description that will follow comes directly from signal detection theory by Macmillan and Creelman (1991).

Signal detection theory accounts for listener responses using a combination of sensitivity and bias measures. Sensitivity is what I was interested in while bias is what some researchers need to take into consideration to recover sensitivity. In the current study, I only report on listeners' sensitivity to Arabic phonemic contrasts. The theory considers sensitivity roughly as detecting a signal against a background noise, and how a listener decides whether the signal is present or absent. Traditionally, experiments presented listeners with signals and non-signals and asked them to respond 'yes' if the signal was present and 'no' if the signal was not. In an AX discrimination task, 'yes' represents a difference (noise) detected by the listener whereas 'no' represents the absence of a difference. To tabulate listeners' responses, I used information from Table 5.9.

	Response: Different (Yes)	Response: Same (No)
Stimuli: Different	Hit	Miss
Stimuli: Same	False Alarm	Correct Rejection

Table 5.9 Information used to tabulate listeners' responses in the AX discrimination task

To explain how the table above was used to tabulate the listeners' responses, I will use an example from the AX discrimination task used in this study. In an Excel sheet, I entered the responses of each listener for each consonant pair using information from Table 5.9. For instance, if a listener judged the pair /qaːn/ - /χaːn/ as 'different', I recorded her response as a 'Hit'. However, if the subject judged this pair as 'same', I recorded her response as a 'Miss'. For a pair such as /qaːn/ - /qaːn/, if the subject judged the pair as 'different', I recorded her response as 'False Alarm'; whereas, if the subject responded 'same', I recorded her response as 'Correct Rejection'.

Taking into account the total number of 'different' trials and the total number of 'same' trials, one could obtain the value in one column from the value in the other. For instance, if the total number of 'different' trials is 30, and the subject has 10 hits, then the subject must have 20 misses. Therefore, when it comes to calculations, only two out of the four numbers in Table 5.9 per row are used to report a subject's performance. The convention is to use the Hit and False Alarm raw numbers.

Following signal detection theory, I used the d' (d-Prime) to statistically measure the difference between the z-transforms of the Hit and False alarm rates with the equation in 5.2.

(5.2)

$$D' = z(H) - z(F)$$

H=Hit, F= False Alarm

The Hit rate is the proportion of 'different' trials to which a subject responded 'different'.[21] The False Alarm rate is the proportion of 'same' trials to which a subject responded 'different'. I got the z-score conversions of the hit and false alarm rates from Table A5.1 in Appendix 5 from Macmillan and Creelman (1991). The d prime value indicates the sensitivity of the listener to the phonemic contrast. A linear model was then used to examine FDHs' discriminability of consonantal contrasts compared to the control group. Listeners' d prime was the dependent variable. The group (FDH or NS) was the independent variable. In order to examine which factors play a role in listeners' discriminability of the consonant contrasts, a linear model was used with listeners' d prime as the dependent variable and LoR (continuous), native language literacy (continuous) and Arabic literacy (categorical) as predictors. When more than one predictor is correlated, problems might arise that pertain to the interpretation of the coefficients; their t-tests and standard errors might be misleading due to the effects of multicollinearity (Mason and Perreault 1991). Multicollinearity does not reduce the reliability or predictive power of the model; it affects each predictor individually and might not give accurate results about which predictor is redundant with the others (Midi *et al.* 2010). The model used was diagnosed for the presence of multicollinearity using the variance inflation factor, *vif()*, that measures the influence of collinearity among the predictors in a regression model. The VIF scores obtained were low (equal to 1), which indicated that it was safe to accurately assess the contribution of the predictors to the model.

5.11.5 Production of Arabic consonants and consonant clusters

A Praat script designed by the author was used to extract all target word productions and their relevant information from Praat and were transferred to Excel files. To examine the accuracy level of the participants, all produced consonants were assigned a value of 0 or 1 depending on whether the production of the sound was correct (target-like) or incorrect (non-target-like). Descriptive statistics based on the percentages of correct/incorrect productions of each consonant averaged across speakers were then provided. To examine which factors play a role in the accuracy of productions, a GLMM was used with LoR (continuous), native language literacy (continuous) and Arabic literacy (categorical) as fixed effects and random intercept of speaker as random effects. The model was detected for multicollinearity using VIF. The scores obtained from the VIF were low indicating low collinearity of the predictors.

[21] A proportion is the total number of a participant's response divided by the total number of trials

With regard to accuracy in producing consonant clusters, a GLMM was used first to examine the difference in modification between onset and coda consonant clusters. First, consonant clusters were assigned a value of 0 or 1 depending on whether the cluster was word initial or word final. Also, productions of the consonant clusters were assigned a value of 0 or 1 depending on whether the production involved modification[22] (yes) or not (no). The GLMM had *modification* (yes or no) as a dependent variable and the type of consonant cluster as fixed effect (predictor). For random effects, speaker was used as random intercept. To examine the effect of social and cognitive factors on pattern of cluster production, onset and coda clusters were analyzed differently. For onset clusters, qualitative analysis was carried out because L1 inventories of some of the FDHs' contained onset clusters while others did not (see Chapter 4). This indicates that L1s of FDHs' are incomparable and will affect their production patterns differently. Therefore, using statistical analysis with factors such as L1, LoR, L1 literacy and L2 literacy might not generate meaningful results and one factor could cancel the significance of other factors (collinearity effects). Therefore, findings from a qualitative analysis will be more reliable and can be indicative of patterns that I can test more thoroughly in future research. As for coda clusters, a GLMM was used to test the effect of LoR, L1 literacy and L2 literacy on modification of consonant clusters. As we have seen in Chapter 4, all FDHs' L1 inventories lacked coda clusters and thus FDHs' were considered comparable in relation to their L1s. All three factors were used as fixed effects with speaker as random intercept for random effects.

5.12 Conclusion

This chapter presented an overview of the study's aims and research questions. This was followed by a description of the hypotheses of the study based on the literature review. The population sample, justification for its selection as well as the social and cognitive characteristics of the FDH participants were provided. A description of the tasks carried out to collect data was presented afterwards. For every task, the stimuli, the task design, and the task procedure were presented in detail. This was followed by a description of the data transcription, labelling, acoustic analyses, and statistical design used to analyze the data and generate results. The following chapter will provide results obtained after statistical analyses were carried out.

[22] This means whether the consonant cluster was maintained or modified by for example epenthesizing a vowel to break it up.

Chapter 6. Foreign Accent Rating of FDHs

6.0 Introduction

This chapter will provide results on foreign accent rating of FDHs. This experiment was caried out to test whether adjustments in NSs' acoustic realizations of consonants and vowels will change based on FDHs' foreign accentedness score. Therefore, it was essential to conduct a rating experiment to get FDHs' foreign accentedness scores.

6.1 Methodology

6.1.1 listeners

The listeners for this task were 10 native Omani speakers (5 males and 5 females, see table 6.1). Their ages ranged between 31 and 40 years at the time they carried out the task. They were all borne in Nizwa and spoke the Nizwa dialect of Omani Arabic. None of them reported having hearing problems.

Listener code	Gender	age
001	Female	33
002	Female	32
003	Female	35
004	Male	31
005	Male	35
006	Male	35
007	Female	38
008	Male	40
009	Female	34
010	Male	37

Table 6.1 Native Omani listeners in the accent rating task

6.1.2 Stimuli

Stimuli for this task were taken from production data collected from the picture-naming task. Ten words from table 5.3 were selected for inclusion in this experiment (/χass/ 'lettuce', / ħabıl/ 'rope', /qalam/ 'pen', /ðˤarf/ 'envelope', /sˤaabuːn/ 'washing liquid', /tˤaːwleh/ 'table', /ʁarʃeh/ 'bottle', /ʕalam/ 'flag', /θoːm/ 'garlic', /ðurah/ 'corn'). The selection of the words was based on the following criterion: the words should contain a complex Arabic consonant word initially. Two native Omani speakers aged 33 and 35 and spoke the Nizwa dialect of Omani Arabic were asked to produce the same words as a control group. The total number of stimuli included in the rating experiment was 240. The stimuli were randomized and each played one time. The listeners had to rate the stimuli which were distributed in three blocks (each including 80 stimuli) on a scale from 1 to 9. The stimuli were extracted from the sound files of the picture-naming task in Praat as whole words. The defined stimuli were windowed by a parabolic function and normalized to 50 dB.

6.1.3 Procedure

The listeners were asked to rate the stimuli they heard on a scale from 1 (not at all native-like) to 9 (completely native-like). They knew that they were going to hear Arabic words produced by native and non-native speakers. They could replay the stimuli up to 100 times and could change their answers if they wished to. When they select an answer, they could move on to the next stimulus by pressing the ‘next’ button. They were offered a break after every 80 stimuli. The listeners spent approximately 30-40 minutes on this experiment.

6.1.4 Data analysis

Descriptive data was generated in R and included mean, median, SD and variance. In order to determine whether there is a difference in rating scores among the two groups (NS and FDHs), I used cumulative link mixed models (CLMM), using the *ordinal* package in R (Christensen 2019). For the dependent variable, I used an ordered factor, the rating response. For the dependent variable or predictor, I used group. For random effects, I used rater and item as intercepts. This was the optimal random effect structure that suited this data. The code used was the following:

model=clmm(response.f~group+(1|rater)+(1|item) , data=data)

6.2 Results of the Foreign Accent RatingTtask

Descriptive results of the foreign accent rating task for all speakers are reported in Table 6.2. The most striking aspect of these results is that the median values of Arabic words in the FDH group were generally low, ranging between 2 and 6, compared to those in the NS group (median=9). Figure 6.1 illustrates the rating response given by raters for the two groups (NS and FDHs) and the percent of total scores for each response. It shows that around 80% of the responses given to NSs fell into the “completely native-like’ category, that is number 9 on the scale. However, the highest percentage in the FDH group went to the “not at all native-like’ ranking on the scale, that is number 1.

Speaker	mean	median	SD	variance
M01	3.67	3	2.59	6.7
M02	3.17	2	2.37	5.65
M03	3.06	2	2.15	4.64
M04	3.33	2.5	2.34	5.49
M05	4.1	3.5	2.33	5.43
M06	5.11	5	2.87	8.28
M07	2.73	2	2.04	4.19
M08	2.57	2	2.02	4.08
M09	3.42	2	2.67	7.13
M10	3.23	2	2.44	5.97
M11	3.09	2	2.34	5.49
M12	4.08	3	2.65	7.04
M13	2.48	2	1.68	5.87
M14	3.98	3	2.65	7.02
M15	4.94	5	2.9	8.46
M16	3.2	2	2.42	5.87
M17	3.73	3	2.49	6.23
M18	3.1	2	2.33	5.46
M19	4.75	4	2.84	8.1
M20	2.7	2	2.08	4.35
M21	4.85	4	3.15	9.92
M22	5.33	6	2.97	8.82
NS	8.58	9	1.36	1.86
NS1	8.65	9	1.06	1.13

Table 6.2 Descriptive statistics of foreign accent rating task

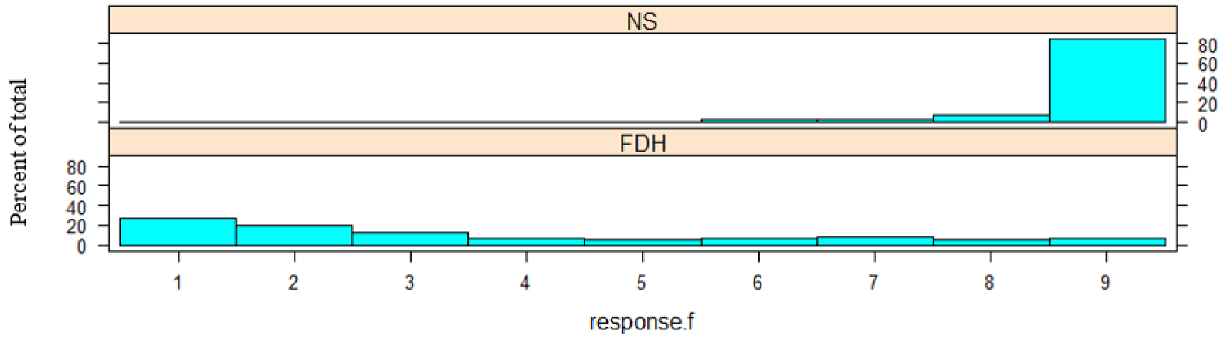


Figure 6.1 Percentages of scores given to NSs and FDHs for each ranking on the scale (1 represents "completely native-like' and 9 represents 'not at all native-like'

Table 6.3 provides descriptive statistics of foreign accent scores based on groups (FDHs and NSs). Median foreign accent score of Arabic words produced by FDHs was lower than that of those produced by NSs. Figure 6.1 illustrates the difference between FDHs and NSs based on median foreign accentedness scores.

Group	mean	median	SD	variance
FDH	3.64	3	2.62	6.86
NS	8.61	9	1.22	1.49

Table 6.3 Descriptive statistics of foreign accent rating task across groups of speakers

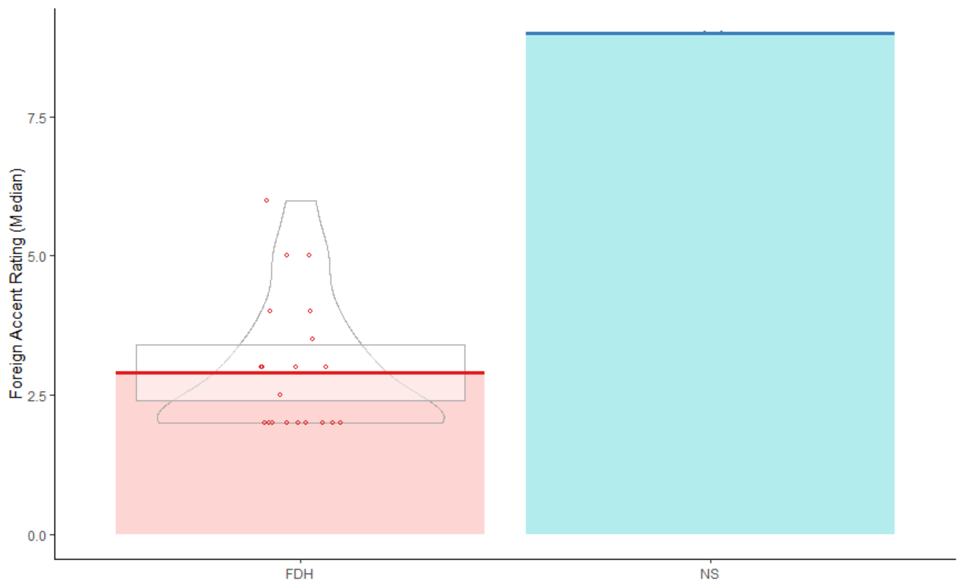


Figure 6.2 Median foreign accent score given to FDHs and the control group

An CLMM revealed that the difference between NSs and FDHs' foreign accentedness was significant as shown in table 6.4. This indicates that FDHs can be considered foreign accented compared to NSs.

	Estimate	Std. Error	z value	Pr(>|z|)
Group: NS	5.31	0.22	22.6	**<2e-16 *****
Random effects	group	variance	SD	
	Rater	1.01	1.006	
	item	0.03	0.19	

Table 6.4 Results of the CLMM model with regard to the effect of group on foreign accent

6.3 Summary and Discussion

This chapter presented a small-scale experiment made to determine any variation in foreign accentedness of FDHs. This was necessary to account for the effect of foreign accentedness of FDHs on hyperarticulation or phonetic changes in FDH-directed speech. The methodology of the experiment including the participants, stimuli and data collection was provided. Data analysis and results were also provided. Results revealed that compared to NSs, FDHs were rated as strongly non-native like. This difference was found to be significant. This conforms to studies that confirmed that second language learners are characterized with a 'foreign accent' despite being exposed to naturalistic input (e.g. Flege 1981; Flege *et al.* 1995; Moyer 2013; Huang 2014). Despite having different LoRs in the Arab world and being exposed to naturalistic input, foreign accentedness scores of FDHs were very similar, indicating their non-native like productions. Since evidence is now established as to the foreign accentedness scores of the FDHs, it would be interesting to find out whether this will be advantageous for this group in terms of the clarity of input they will receive. Moyer (2009) suggested that having a foreign accent can lead NSs to clarifying their speech in interaction with foreigners. Thus, part of the analyses that will be carried out in the next Chapter includes an examination of the influence of FDHs' foreign accentedness scores on hyperarticulation in NSs' productions.

Chapter 7. Realization of Consonants, Consonant Clusters, and Vowels in FDH-Directed Speech

7.0 Introduction

This Chapter will provide results on the realization of consonants, consonant clusters, and vowels in FDH-Directed Speech. Section 7.1 will provide results on the realization of complex Arabic consonants based on auditory analysis. This analysis was based on whether consonants in FDH-directed speech sounded native/non-native-like by comparing them to ADS. The term 'native-like' analysis in this study refers to whether the sounds produced by the NSs when addressing FDHs or the NS sounded native-like or not to the researcher. As mentioned in chapter 6, the researched speaks the same dialect as the NSs in this study. Section 7.2 presents acoustic profiling of the consonants under examination. This analysis was carried out as a reliability check for the auditory analysis. Section 7.3 provides results regarding the realization of FDH-directed speech complex onsets and codas. Results on the frequency with which the NSs either delete or retain the short unstressed vowel in onset consonant clusters based on who their interlocutor is will be provided. In addition, results on the frequency with which the NSs would modify coda consonant clusters based on who the interlocutor is will be reported. Results on the effect of other factors (e.g. condition order on any significant modifications in FDH-directed speech or ADS) will be provided. Furthermore, results on the effect of LoR on any adaptations observed in FDH-directed speech will be presented.

Section 7.4 will provide results regarding the realization of the plain/emphatic consonants (/s/ vs. /sˤ/ and /t/ vs. /tˤ/) in FDH-directed speech based on acoustic analyses. To rule out any effect of other factors on significant adaptations observed in the metrics examined, results on the role of prosodic boundary, condition order and word stress will be provided where relevant. Furthermore, results regarding the effect of foreign accentedness, LoR and religion on all metrics examined will be provided.

Section 7.5 will provide results on the realization of vowels in FDH-directed speech. Results on vowel space size and other vowel properties including F1, F2, intensity, duration and *f*0 will be provided. Also, for each metric examined, results on the effect of other factors (e.g. prosodic boundary and condition order) will be presented to confirm that any significant changes in the above-mentioned vowel properties is not a consequence of other unintended variations in speaking style or condition order. The role of foreign accentedness, LoR and religion on changes in all metrics examined will also be provided.

Before presenting the results in the following sections, it is important to point out that after impressionistic analysis was carried out, target words that appeared as focus words were the ones used for analyses. In fact, 98% of the target words in the data (both FDH-directed speech and ADS) were found to appear as focus words.

7.1 Realization of Complex Arabic Consonants: Auditory Analysis

Auditory analysis of complex Arabic consonants was set to test the following hypotheses:

H1. NSs will simplify complex consonants and produce sounds that are familiar to FDHs (e.g. /ʁ/→/g/; /q/→/k/).

H4. Simplification and hyperarticulation in FDH-directed speech will vary based on FDHs' foreign accentedness score, LoR and religion.

Table 6.1 shows the number of tokens for each target consonant by interlocutor and the proportion of target-like productions of each consonant. Descriptive statistics provided evidence that NSs in this study did not simplify their consonants but rather produced all consonants in the native form when addressing FDHs. It is interesting to observe that the percentage of native-like productions in both FDH-directed speech and ADS were very similar. FDH-directed speech had 94.73% native-like productions, while ADS had 92.3% native-like productions.

There was variation in the production of /sˤ/ specifically in the target word /sˤaħan/ that resulted in non-native like productions (around 5% in FDH-directed speech and 7% in ADS) due to de-emphasis. Three NS participants deemphasized the emphatic consonant /sˤ/ and produced the word as [saħan] throughout their interaction with both the FDH and the NS. This modification of the emphatic fricative in this particular lexical item appeared to be speaker-specific.

Consonant	FDSH-directed speech		ADS	
	Tokens	% Native-like	Tokens	% Native-like
/ð/	76	100	39	100
/ðˤ/	66	100	28	100
/θ/	77	100	67	100
/ʕ/	77	100	40	100
/ħ/	99	100	59	100
/ʁ/	66	100	52	100
/χ/	57	100	47	100
/q/	63	100	43	100
/tˤ/	79	100	62	100
/sˤ/	76	94.73	52	92.307
Total	736	94.73	489	92.3

Table 7.1 Proportion of native-like productions of the target consonants by interlocutor

7.2 Realization of Complex Arabic Consonants: Acoustic Profiling

To support auditory analysis, this section provides acoustic profiling of all consonants under examination. It provides a description of the way the consonants under investigation were realized in FDH-directed speech and ADS as evident on spectrograms.

7.2.1 The alveolar emphatic consonants

7.2.1.1 /sˤ/

Figures 7.1 and 7.2 show spectrograms of /sˤ/ as it appeared in FDH-directed speech and ADS. In both figures, /sˤ/ exhibits frication noise in higher frequencies. Linguists working on Arabic have reported no correlation between the bottom clear-cut off frequency in spectrograms and the presence or absence of emphasis (e.g. Card 1983 Norlin 1987; Al-Khairy 2005). Spectrograms of /s/ as produced in FDH-directed speech and ADS are shown in Figures 7.3 and 7.4. Noticeable differences between /sˤ/ and /s/ appear in frequency transitions at the start of voicing of the following vowel. Lowering of F2 and rising of F1 that results in an approximation of these two frequencies during the vowel production is clearly illustrated in the spectrograms of /sˤ/ in both speech styles.

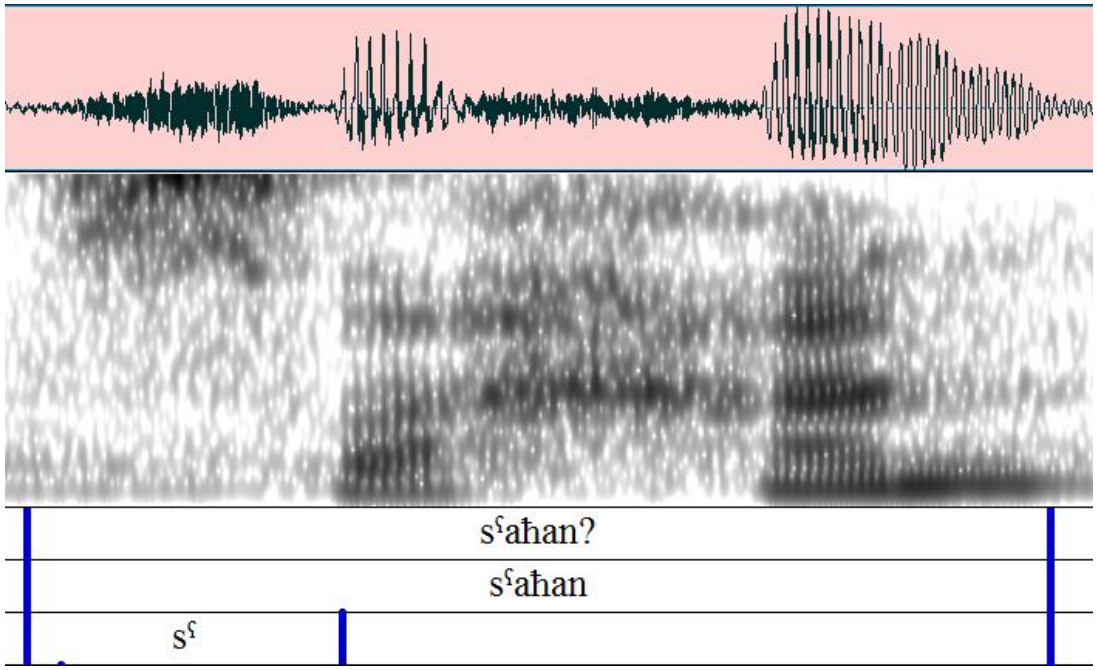


Figure 7.1 Spectrogram and waveform of the emphatic fricative /sˤ/ in the word *sˤaħan* in FDH-directed speech

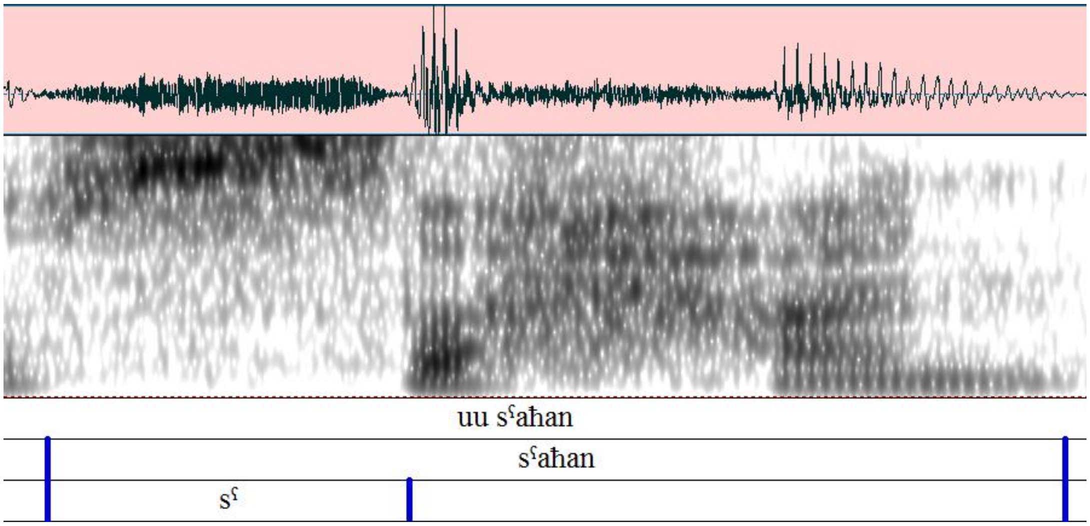


Figure 7.2 Spectrogram and waveform of the emphatic fricative /sˤ/ in the word *sˤaħan* in ADS

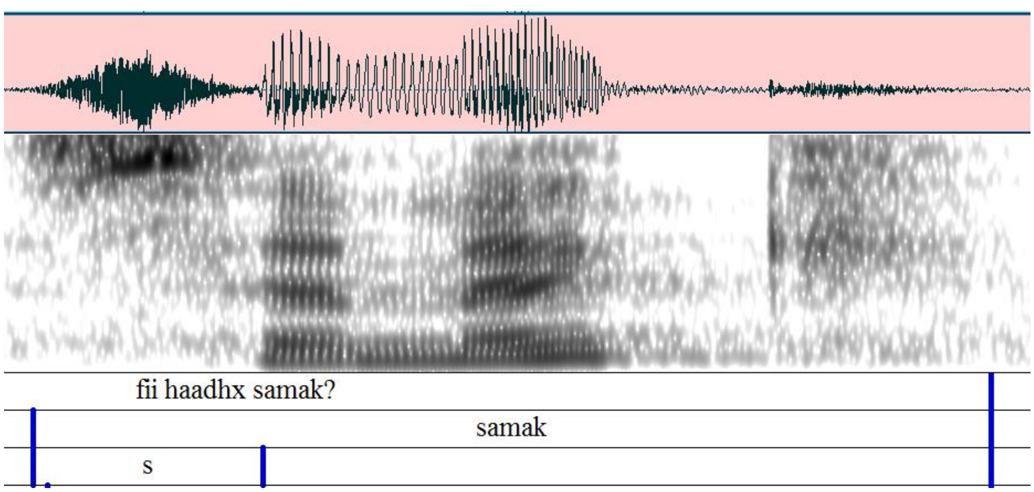


Figure 7.3 Spectrogram and waveform of the alveolar fricative /s/ in the word *samak* in FDH-directed speech

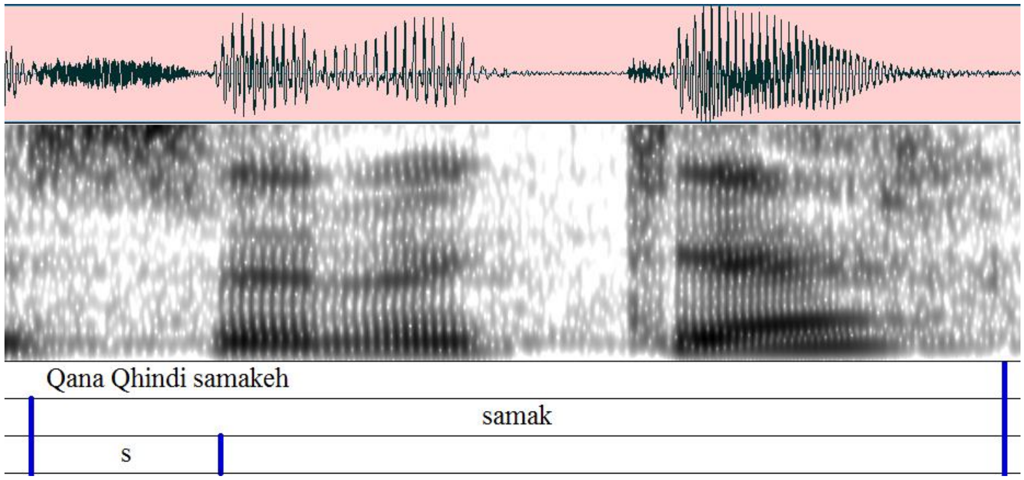


Figure 7.4 Spectrogram and waveform of the alveolar fricative /s/ in the word *samak* in ADS

7.2.1.2 /tˤ/

Figures 7.5 and 7.6 show two examples of /tˤ/ from ADS and FDH-directed speech as it appeared in spectrograms. Similar to what has been reported about this sound in the literature (e.g. Al-Ani 1970; Card 1983; Zawaydeh 1999; Bin Muqbil 2006), both figures show a burst followed by short aspiration and lowering of F2 of the following vowel. Compared to /tˤ/, /t/ exhibits strong aspiration following the bursts in both ADS and FDH-directed speech in line with previous studies (e.g. Al-Ani 1970; Giannini and Pettorino 1982; Al-Nuzaili 1993) (Figures 7.7 and 7.8).

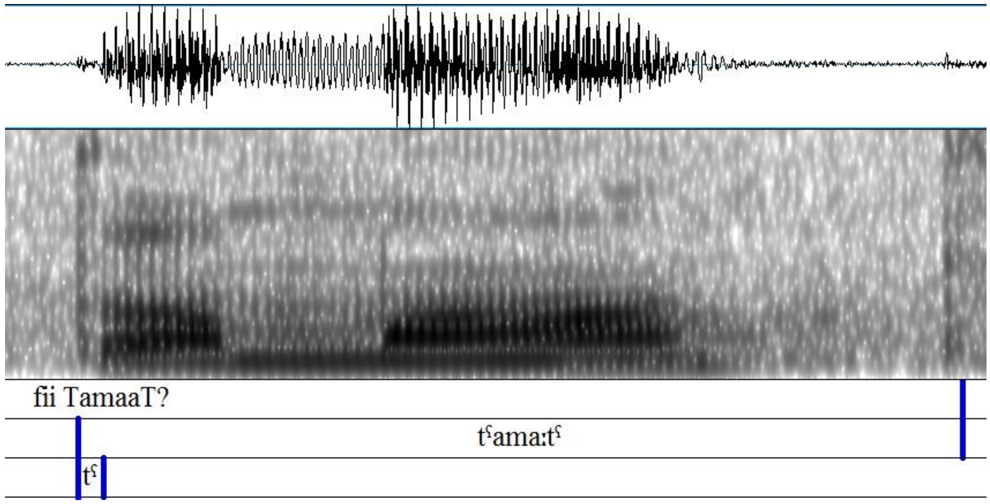


Figure 7.5 Spectrogram and waveform of the emphatic stop /tˤ/ in the word *tˤamatˤ* in FDH-directed

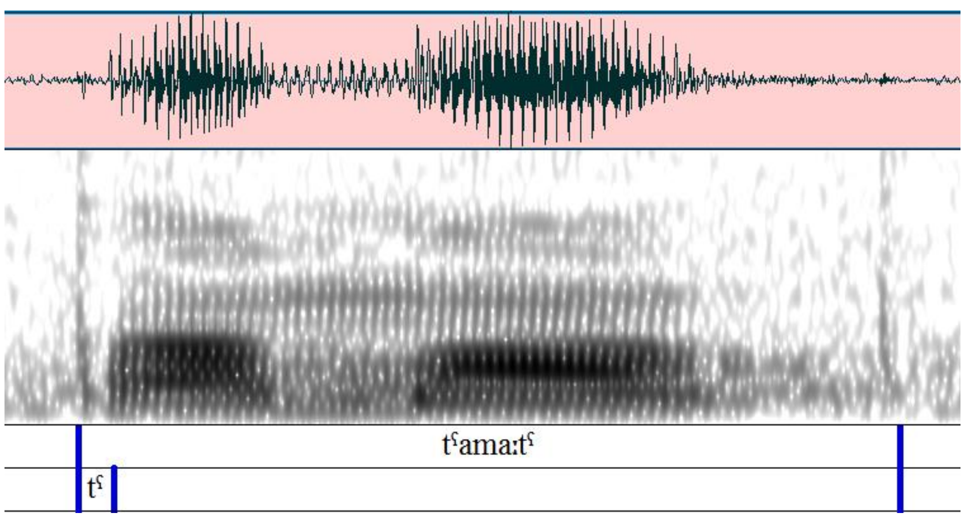


Figure 7.6 Spectrogram and waveform of the emphatic stop /tˤ/ in the word *tˤamatˤ* in ADS

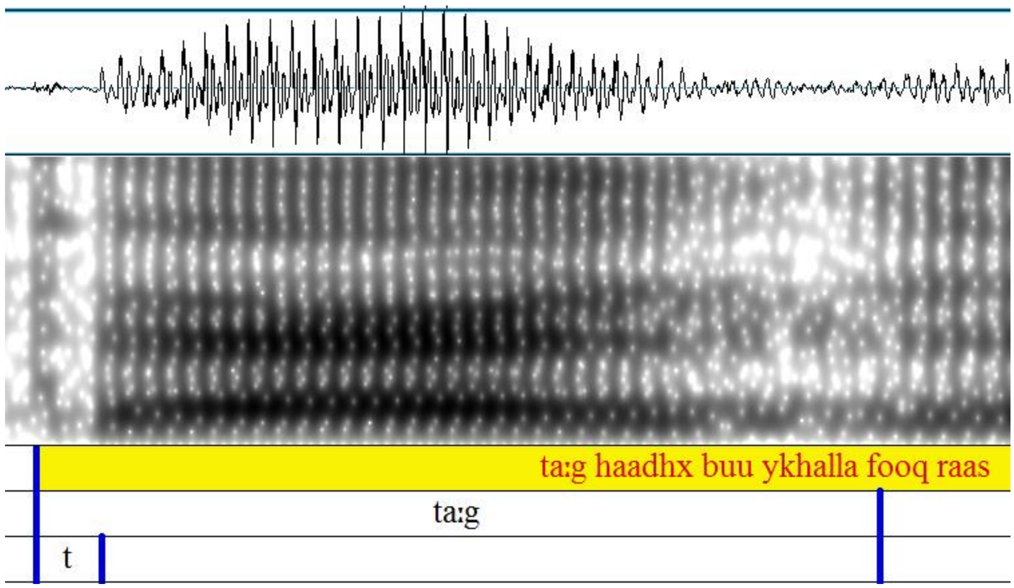


Figure 7.7 Spectrogram and waveform of the alveolar stop /t/ in the word *ta:g* in FDH-directed speech

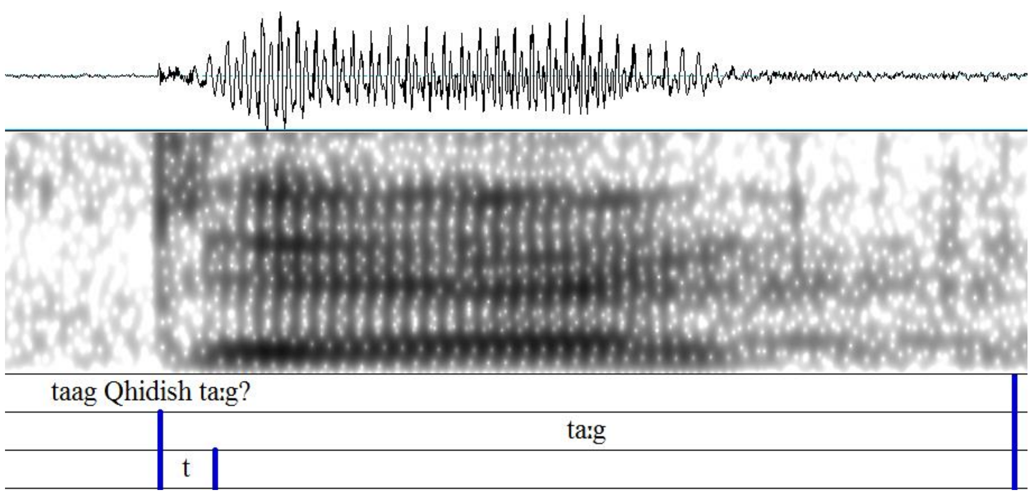


Figure 7.8 Spectrogram and waveform of the alveolar stop /t/ in the word *ta:g* in ADS

7.2.2 The pharyngeals

7.2.2.1 /ʕ/

/ʕ/ was found to appear in three different forms in spectrograms: an approximant, a fricative, and a stop-like. All three variants for this variable have been reported in the literature (Al-Ani 1970; Heselwood 2007; Alotaibi and Muhammad 2010). Table 7.2 shows the proportions of the different realizations of /ʕ/ in FDH-directed speech and ADS. In both speech styles, chances of /ʕ/ appearing as an approximant were more than those of it appearing as a fricative or a stop. Only in FDH-directed speech was /ʕ/ realized as stop-like and that was in 4.27% of the total productions. The three realizations of /ʕ/ will be discussed in detail in the following lines.

	FDH-directed speech	ADS
Approximant	59 (72.64%)	26 (65%)
Fricative	15 (23%)	14 (35%)
Stop	5 (4.27%)	0

Table 7.2 The number of tokens of each realization of the pharyngeal /ʕ/ within register and their proportions between brackets

- Approximant

Figures 7.9 and 7.10 show examples of the pharyngeal /ʕ/ appearing as an approximant in FDH-directed speech and ADS. Classifying /ʕ/ as an approximant or glide-like is the most common in the literature according to Heselwood (2007). When /ʕ/ was realized as an approximant, it appeared in spectrograms with slightly lower energy compared to neighboring vowels. Energy in the higher frequencies was lost. Also, a number of researchers have established that an approximant /ʕ/ may be detected because of a lowering in F2 and a rising in F1 which indicates some degree of narrowing in the pharynx as indicated by the red arrows (Delattre 1968; Fant 1973; Laver 1980; Butcher and Ahmad 1987).

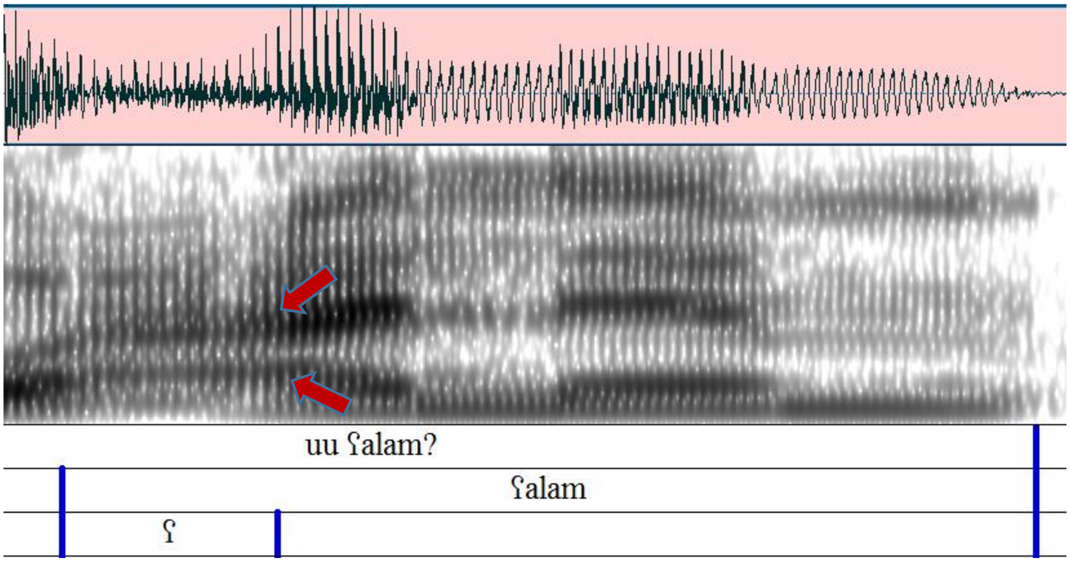


Figure 7.9 Spectrogram and waveform of the pharyngeal /ʕ/ realized as an approximant in the word *ʕalam* in FDH-directed speech

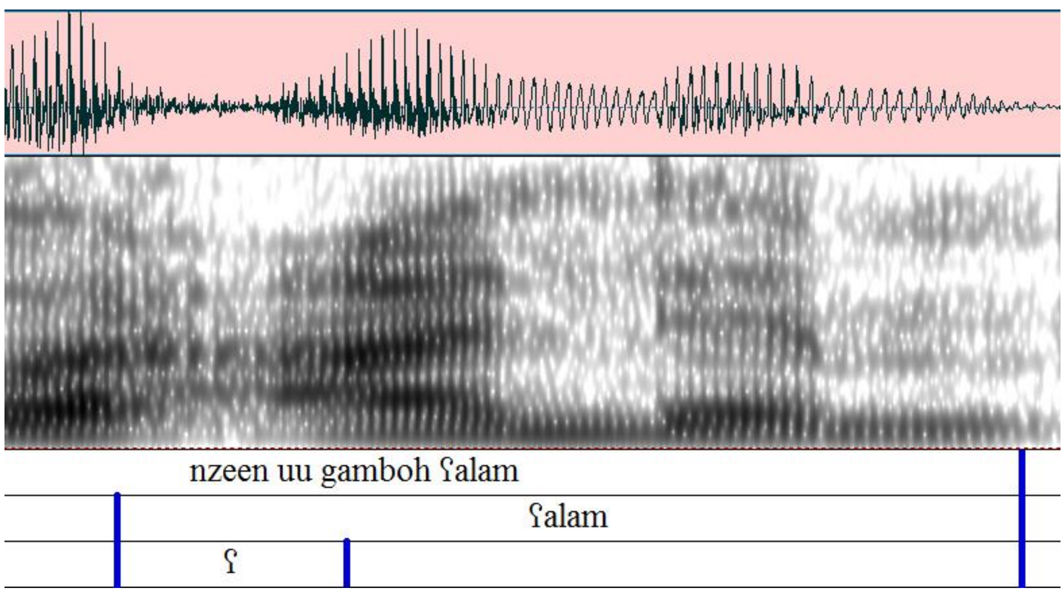


Figure 7.10 Spectrogram and waveform of the pharyngeal /ʕ/ realized as an approximant in the word *ʕalam* in ADS

Figure 7.11 shows other examples where the pharyngeal may be classified as an approximant or vowel-like. In these examples and in similar other cases, a boundary between the pharyngeal and the following vowel was only discernible through the lower amplitude especially in the higher frequencies. The consonant could also be detected based on an approximation between F1 and F2, which could be detected at the onset of the labelled boundary in the spectrogram provided. /ʕ/ was most likely realized as a pharyngeal blended with the following vowel when it appeared in words that were at a phrase initial position like the one in the example or when preceded by a long pause. Phrase initial positions normally attract stop-like /ʕ/ realizations and not vowel-like realizations (Heselwood 2007).

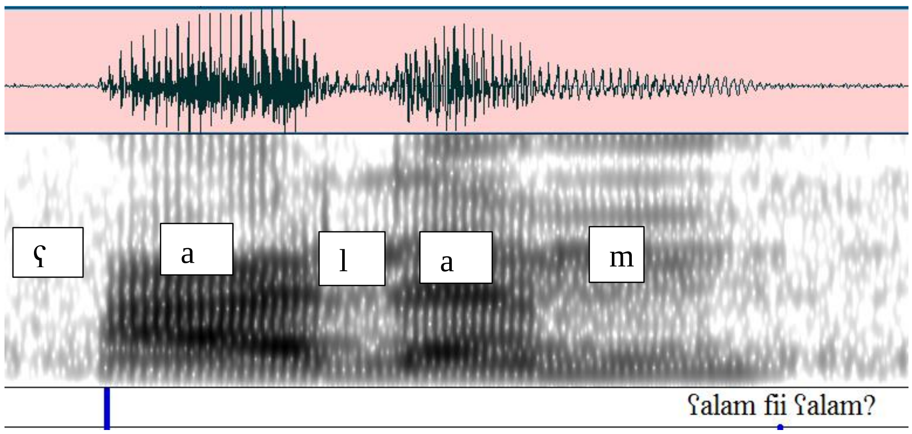


Figure 7.11 Spectrogram and waveform of the pharyngeal /ʕ/ realized as an approximant with no clear-cut boundary from the following vowel in the word *ʕalam* in FDH-directed speech

- Fricative

When /ʕ/ was realized as a fricative (See, for example, Figures 7.12 and 7.13), some turbulence was visible in the waveform and some random energy or friction appeared in spectrograms. Several researchers have classified /ʕ/ as a fricative (e.g. Ghazeli 1977; Laradi 1983; Abu-Haidar 1991; Holes 2004; Alotaibi and Muhammad 2010). The number of cases in which /ʕ/ appeared as a fricative in this study was comparable across ADS and FDH-directed speech as Table 6.2 demonstrates.

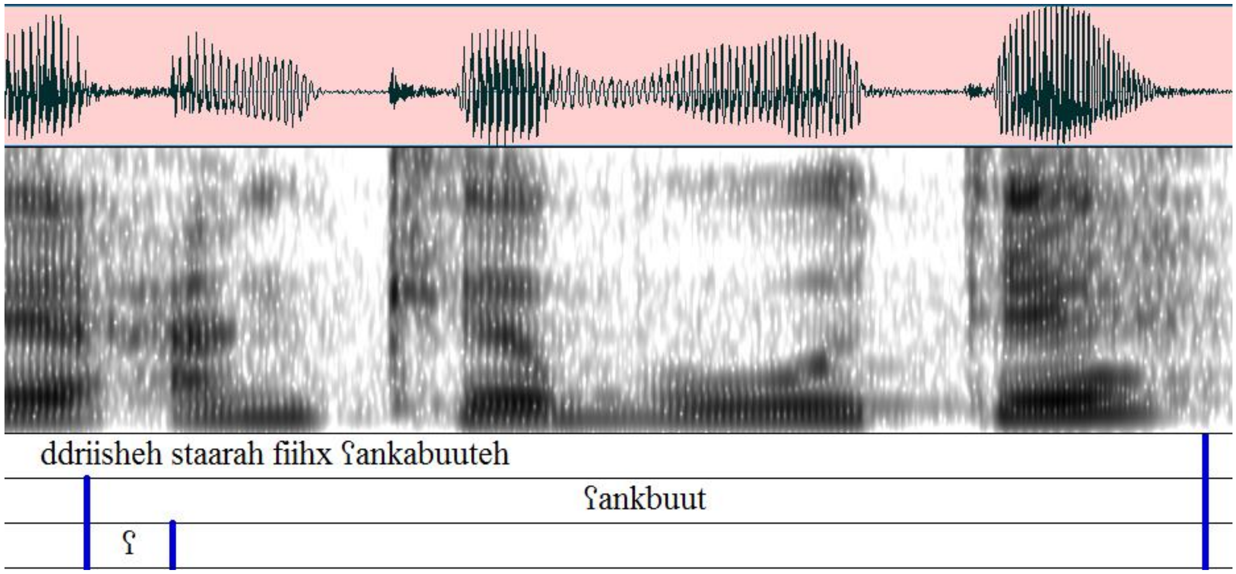


Figure 7.12 Spectrogram and waveform of the pharyngeal /ʕ/ realized as a fricative in the word *ʕankbuːt* in FDH-directed speech

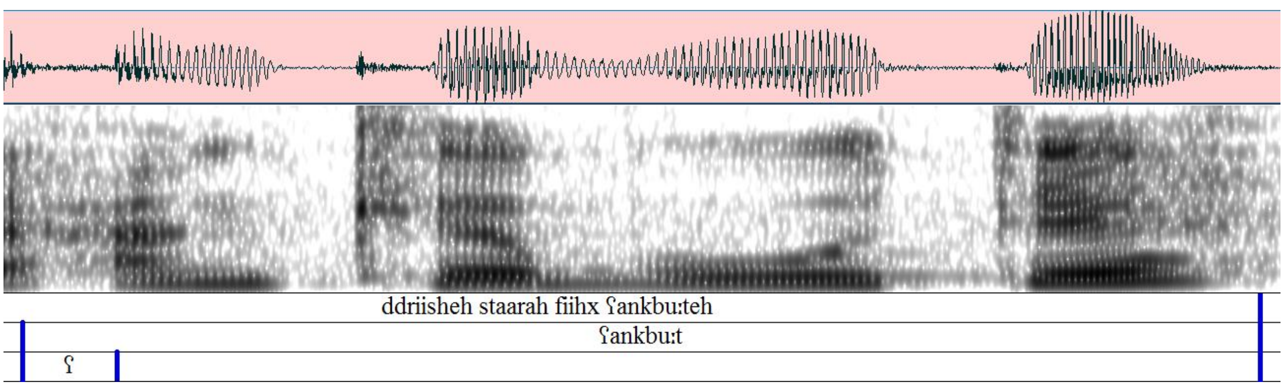


Figure 7.13 Spectrogram and waveform of the pharyngeal /ʕ/ realized as a fricative in the word *ʕankbuːt* in ADS

- Stop-like

In very few cases and only in FDH-directed speech, /ʕ/ was realized as a voiceless stop. Figure 7.14 shows /ʕ/ starting with a burst followed by a silent gap and then some creak and formants which indicates the onset of the following vowel. Stopped realizations are common among highly conservative dialects and are more likely to appear in constrained phonological contexts, with intervocalic positions the least prone to complete obstruction (Heselwood 2007). Indeed, in all five cases where /ʕ/ appeared as a stop, the target words containing /ʕ/ were in an initial phrase position or preceded by a long pause. One explanation for the absence of a stop realization of /ʕ/ in ADS could be that the number of /ʕ/ tokens in FDH-directed speech was far more than that in ADS, which means that there were far more chances for /ʕ/ to be realized as a stop in FDH-directed speech than there were in ADS. Another possibility is that /ʕ/, according to Laufer and Condax (1979), can undergo complete closure in careful and slow speech, and in this case, FDH-directed speech could be considered a good example of a careful and slow speech, based on research mentioned in Chapter 3 (but the current study did not find FDH-directed speech to be slower, at least based on segment duration as will be shown in sections 7.4 and 7.5). Whether and how a stop realization of /ʕ/ is a result of slow speech, future research will have to further investigate this.

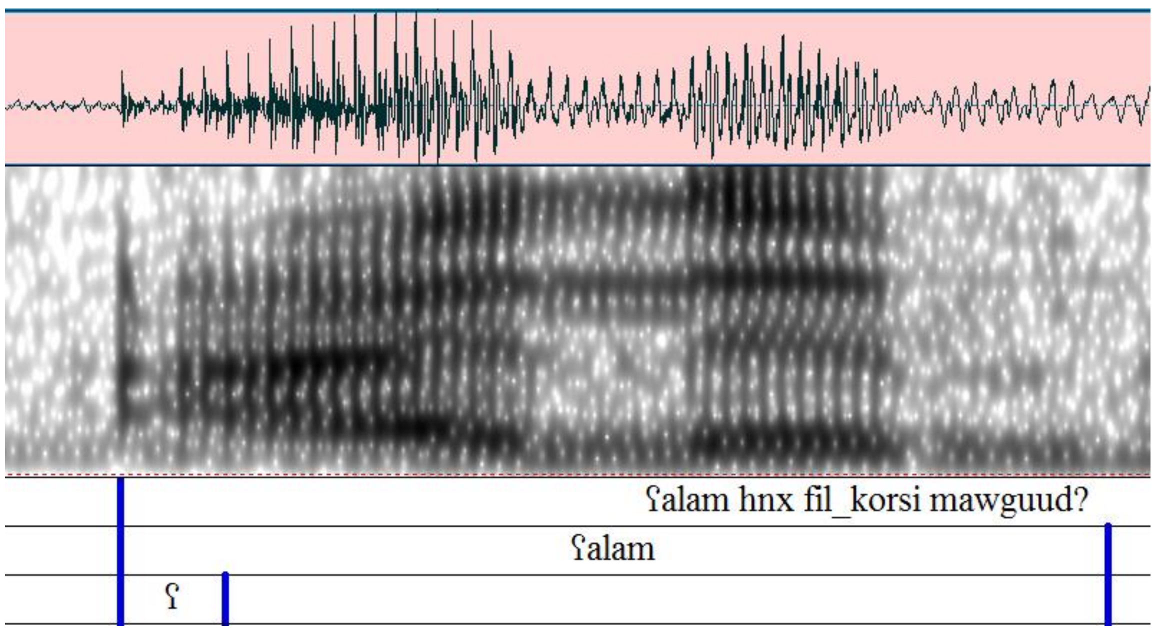


Figure 7.14 Spectrogram and waveform of the pharyngeal /ʕ/ realized as a voiceless stop in the word *ʕalam* in FDH-directed speech

7.2.2.2 /ħ/

/ħ/ appeared in spectrograms of both ADS and FDH-directed speech exhibiting pseudo-formants at different frequencies below 4000 Hz as reported in the literature (e.g. Al-Ani 1970; McCarthy 1994) (See Figures 7.15 and 716). Although some studies reported that the pharyngeals can have vowel-like formants (e.g. Al-Ani 1970), the patterning of /ħ/ in the data of this study appeared more as fricative-like in both ADS and FDH-directed speech.

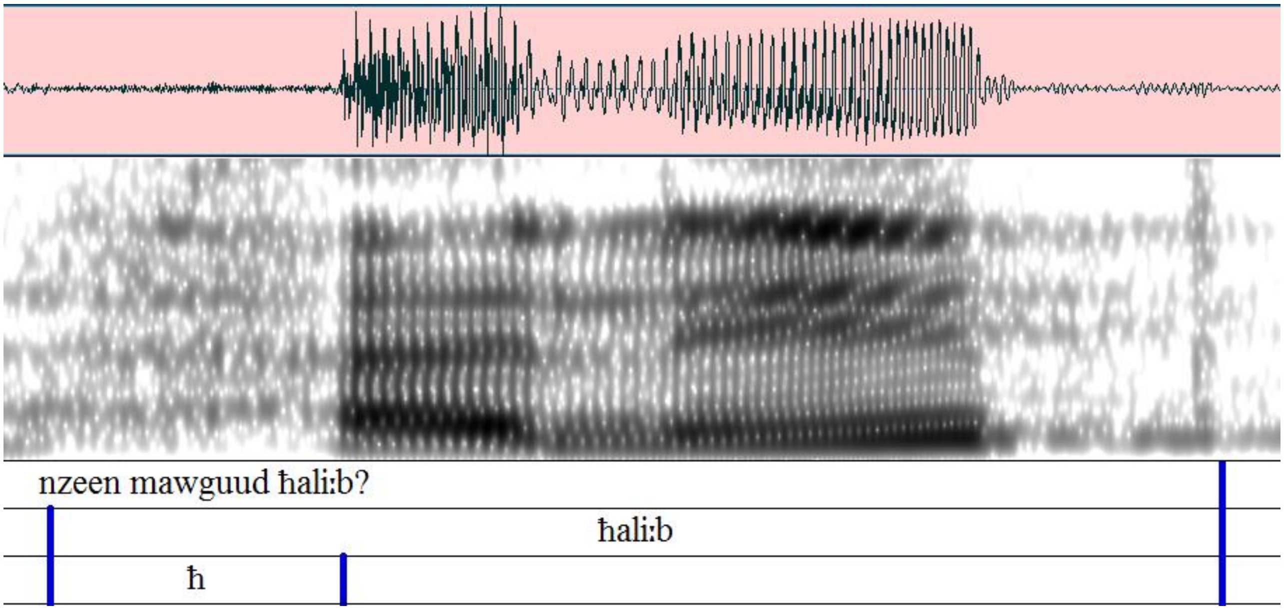


Figure 7.15 Spectrogram and waveform of the pharyngeal /ħ/ realized as a fricative in the word *ħali:b* in FDH-directed speech

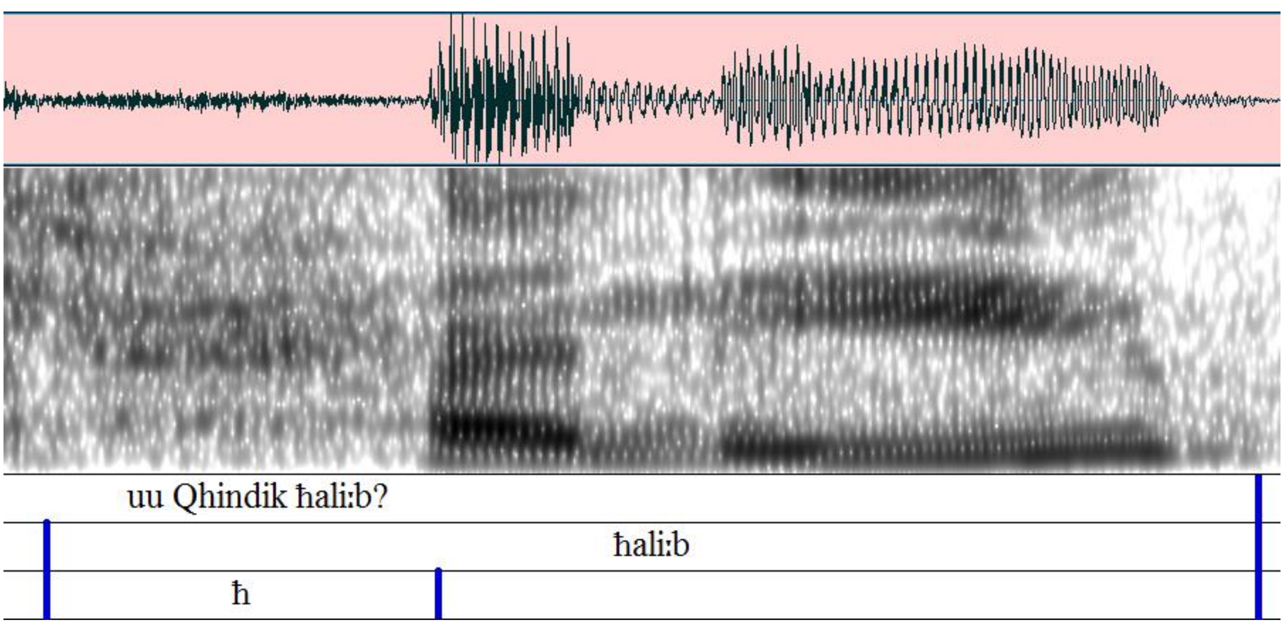


Figure 7.16 Spectrogram and waveform of the pharyngeal /ħ/ realized as a fricative in the word *ħali:b* in ADS

7.2.3 The uvulars

7.2.3.1 /ʁ/

Although the uvular /ʁ/ has generally been described as having vowel-like formants (Al-Ani 1970) or as a voiced trill (McCarthy 1994), it was detected exhibiting some other acoustic features in spectrograms of the current study. Accordingly, it is characterized into three classes: approximant, fricative and stop-like. Table 7.3 shows the proportions of /ʁ/ with regard to its spectrographic classification within each speech style.

	FDH-directed speech	ADS
Vowel-like	28 (23.72%)	12 (23%)
Fricative	86 (72.88%)	40 (73%)
Stop-like	4 (3.49%)	2 (3.85%)

Table 7.3 The number of tokens of each realization of the uvular /ʁ/ within each interlocutor group and their proportions between brackets

- Vowel-like

Figures 7.17 and 7.18 show /ʁ/ featuring vowel-like formants in FDH-directed speech and ADS. /ʁ/ appears with slightly lower energy compared to neighboring vowels just like /ʕ/, with some weak noise throughout its duration indicating a slightly fricative manner of articulation (Al-Ani 1970). We can also notice the lowering of F2 and rising of F3.

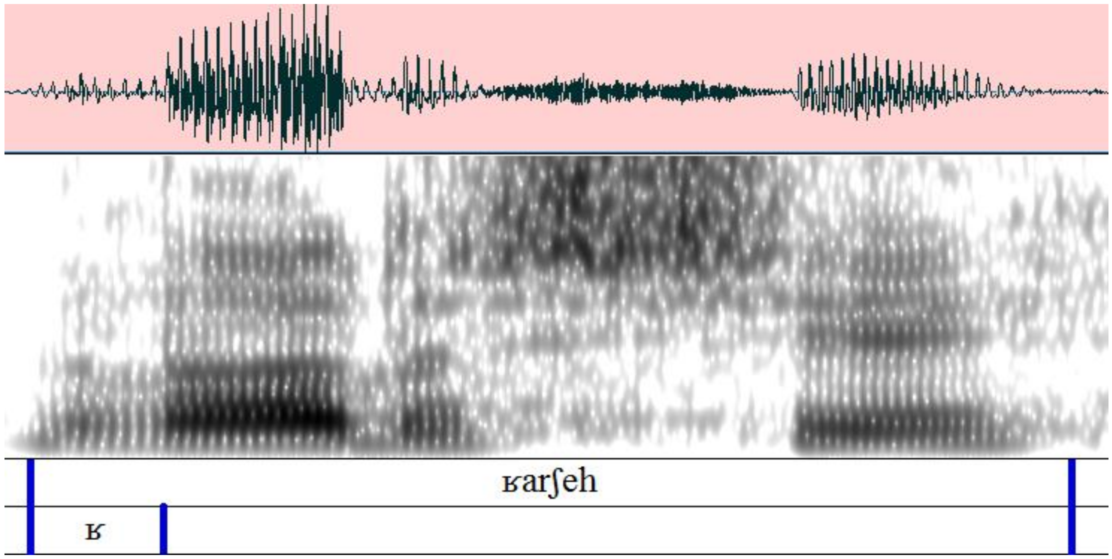


Figure 7.17 Spectrogram and waveform of the uvular /ʁ/ realized as an approximant in the word *ʁassa:leh* in FDH-directed speech

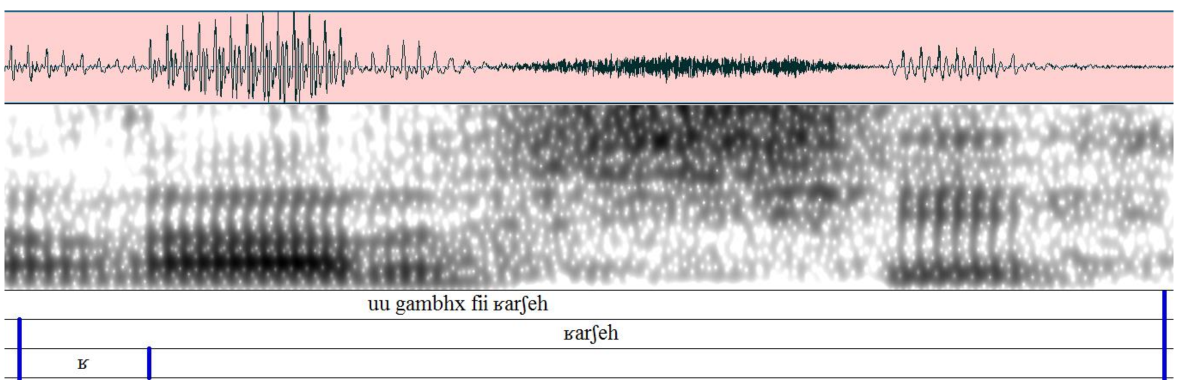


Figure 7.18 Spectrogram and waveform of the voiced uvular /ʁ/ realized as an approximant in the word *ʁarʃeh* in ADS

- Fricative

In most cases, /ʁ/ appeared in spectrograms as fricative-like. Figures 7.19 and 7.20 illustrate examples of a fricative /ʁ/ with random noise throughout the segment's duration in FDH-directed speech and ADS. Classifying the voiced uvular as a fricative is in line with that of Bin-Muqbil (2006) who found that Arabic uvular continuants involve considerable airflow turbulence in their production pointing to their fricative-like acoustic qualities.

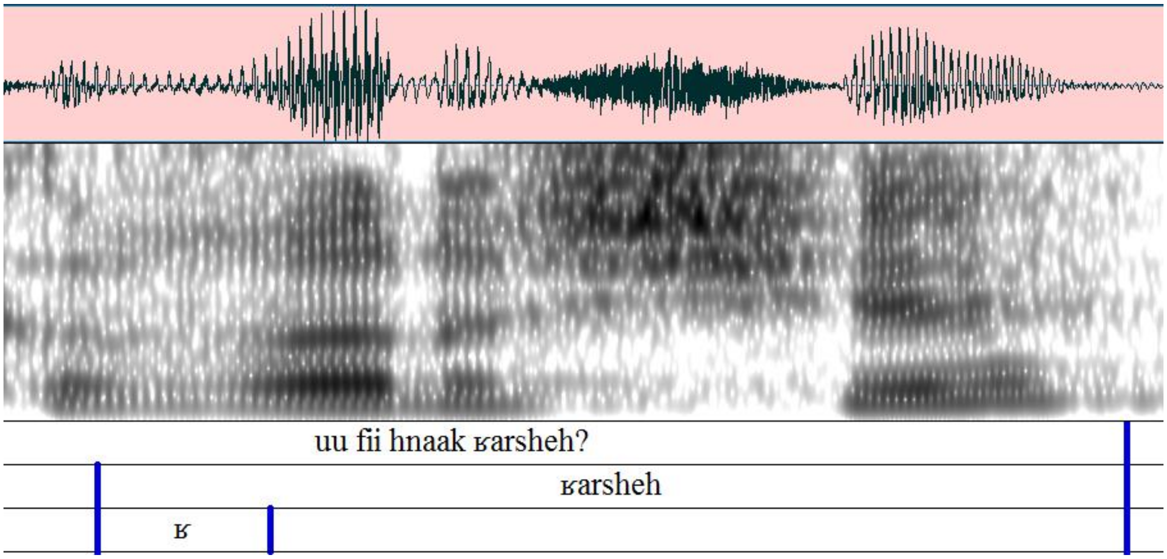


Figure 7.19 Spectrogram and waveform of the voiced uvular /ʁ/ realized as a fricative in the word *ʁarʃeh* FDH-directed speech

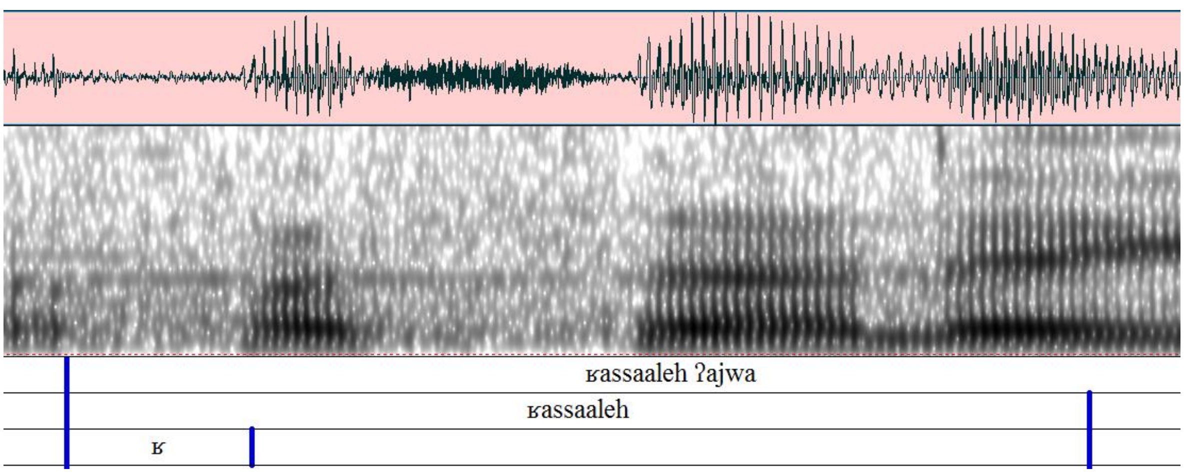


Figure 7.20 Spectrogram and waveform of the voiced uvular /ʁ/ realized as a fricative in the word *ʁarʃeh* in ADS

➢ Stop-like

In very few cases, /ʁ/ appeared as a voiced stop or featuring a stop-like realization as Figures 7.21 and 7.22 show. There is pre voicing followed by a clear burst followed by a gap with some friction leading to the onset of the following vowel. The frequency at which each realization occurred was approximately similar across speech styles given the difference in the number of tokens. To my knowledge, the stop realization of /ʁ/ that occurred in both FDH-directed speech and ADS has not been reported in the literature. It could be a result of spontaneous speech or careful speech.

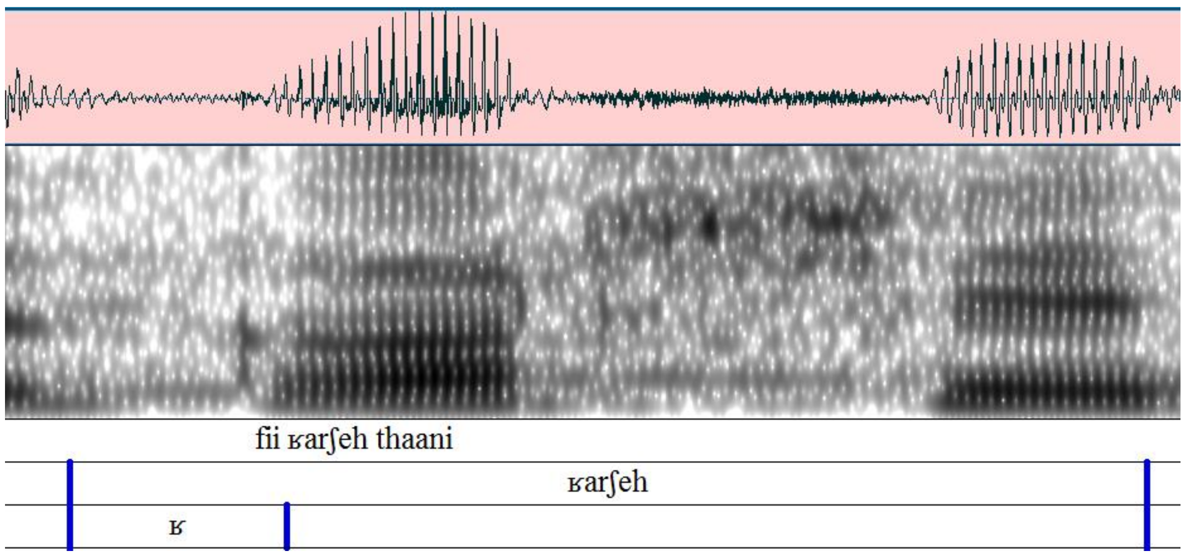


Figure 7.21 Spectrogram and waveform of the voiced uvular /ʁ/ realized as a voiced stop in the word *ʁarʃeh* in FDH-directed speech

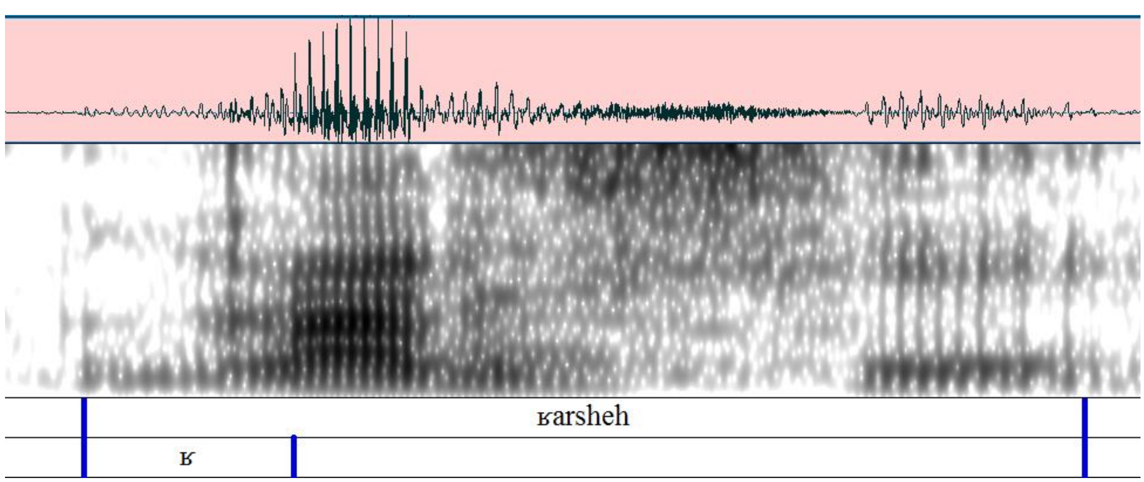


Figure 7.22 Spectrogram and waveform of the voiced uvular /ʁ/ realized as a voiced stop in the word *ʁarʃeh* in ADS

7.2.3.2 /q/

Figures 7.23 and 7.24 show the sound /q/ produced as a voiceless unaspirated stop in both FDH-directed speech and ADS. This description is in line with previous reports on the acoustic properties of this sound (e.g. Al-Ani 1970; Ghazeli 1977; Bin Muqbil 2006; McCarthy 1994). Lowering of F2 of the following vowel is another property reported about the effect of /q/ on neighboring vowels (Al-Ani 1970; Giannini and Pettorino 1982; Bin Muqbil 2006). FDH-directed speech and ADS spectrograms of this sound show a similar pattern.

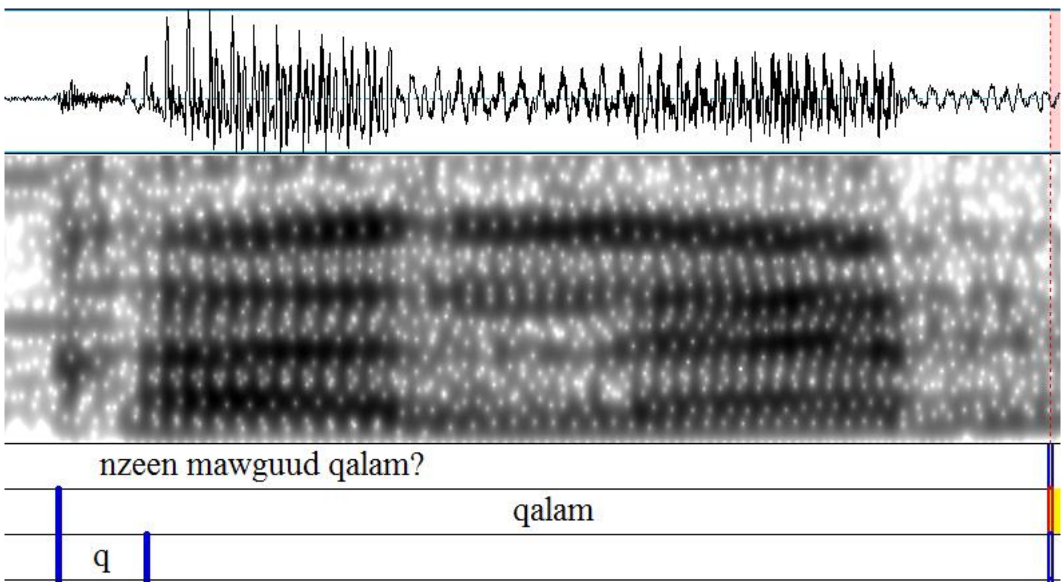


Figure 7.23 Spectrogram and waveform of the voiceless uvular /q/ in the word *qalam* in FDH-directed speech

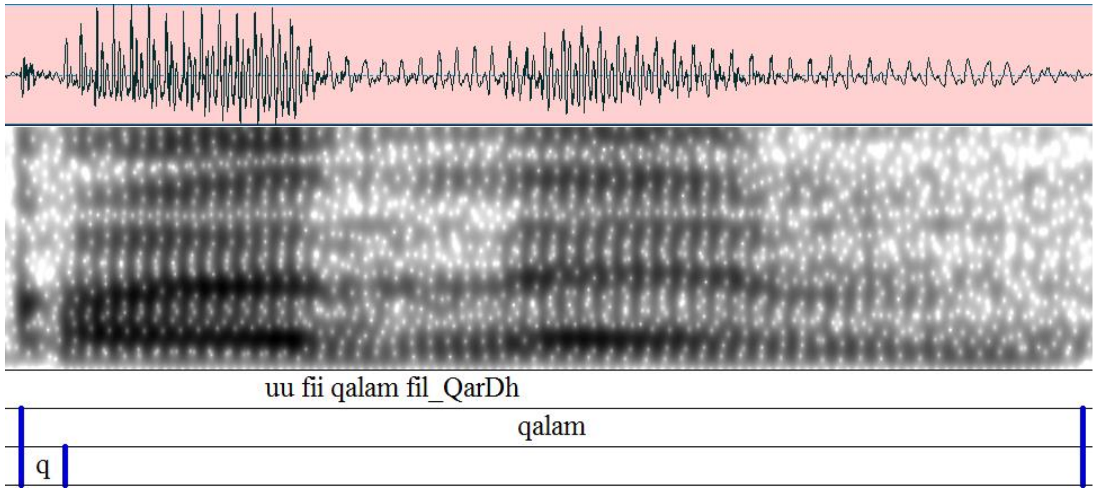


Figure 7.24 Spectrogram and waveform of the voiceless uvular /q/ in the word *qalam* in ADS

7.2.3.3 /χ/

/χ/ appeared in spectrograms with irregular noise during its duration in both FDH-directed speech and ASD as described by Al-Ani (1970) and Ghazeli (1977) (Figures 7.25 and 7.26).

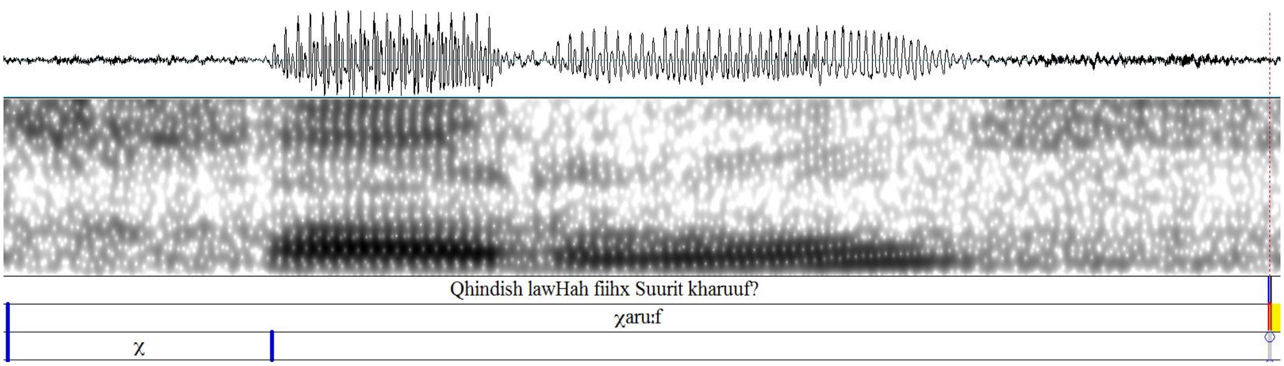


Figure 7.25 Spectrogram and waveform of the voiceless uvular /χ/ in the word *χaru:f* in FDH-directed speech

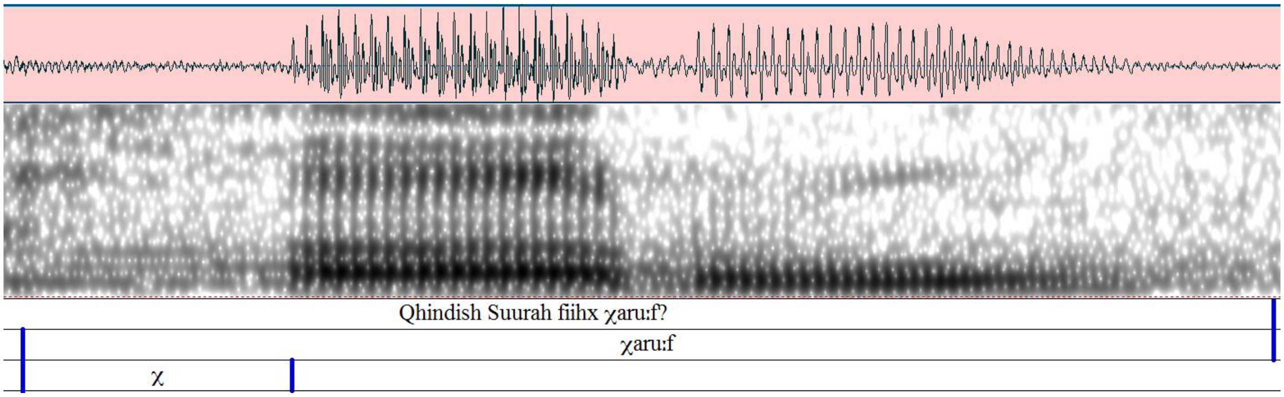


Figure 7.26 Spectrogram and waveform of the voiceless uvular /χ/ in the word *χaru:f* in ADS

7.2.4 The dentals /θ/, /ð/ and /ðˤ/

The dentals /θ/, /ð/ and /ðˤ/ were also found to exhibit variation in their realizations by NSs. They did not merely appear as fricatives, the only description reported on these segments in the Arabic literature, but also as stop-like. Table 7.4 provides proportions of the two realizations of each of these segments in FDH-directed speech and ADS.

Dentals	FDH-directed speech		ADS	
	Fricative	Stop	Fricative	Stop
/θ/	66 (83.33%)	11 (16.77%)	54 (80.6%)	13 (19.4%)
/ð/	62 (80%)	14 (20%)	30 (77%)	9 (23%)
/ðˤ/	24 (80.85%)	4 (19.15%)	24 (85.71%)	4 (14.28%)

Table 7.4 The number of tokens of each realization of the three dentals within speech style and their proportions between brackets

I would classify a stop-like realization of the dental sounds as non-native-like and non-target-like. Although the realization of dental sounds as dental stops is common in urban dialects, in Oman this realization does not exist and would be considered non-native.

- Fricative

As described in Chapter 4, Arab linguists describe /θ/, /ð/ and /ðˤ/ as interdental fricatives (Al-Ani 1970). Figures 7.27 and 7.28 show examples of /θ/ realized as a fricative in both FDH-directed speech and ADS. In line with Al-Ani (1970), /θ/ exhibited random noise throughout the entire segment. Figures 7.29 and 7.30 show examples of /ð/ realized as a fricative in both speech styles. The spectrograms show weaker overall energy compared to those of /θ/ exactly as has been described by Al-Ani (1970). Figures 7.31 and 7.32 illustrate examples of /ðˤ/ realized as a fricative. It can be detected in the spectrogram with weak overall energy similar to that of /ð/. However, rising of F1 and lowering of F2 of the following vowel makes it distinct from /ð/. Spectrograms of /ð/ and /ðˤ/ are also distinguished from those of /θ/ by the voicing bar.

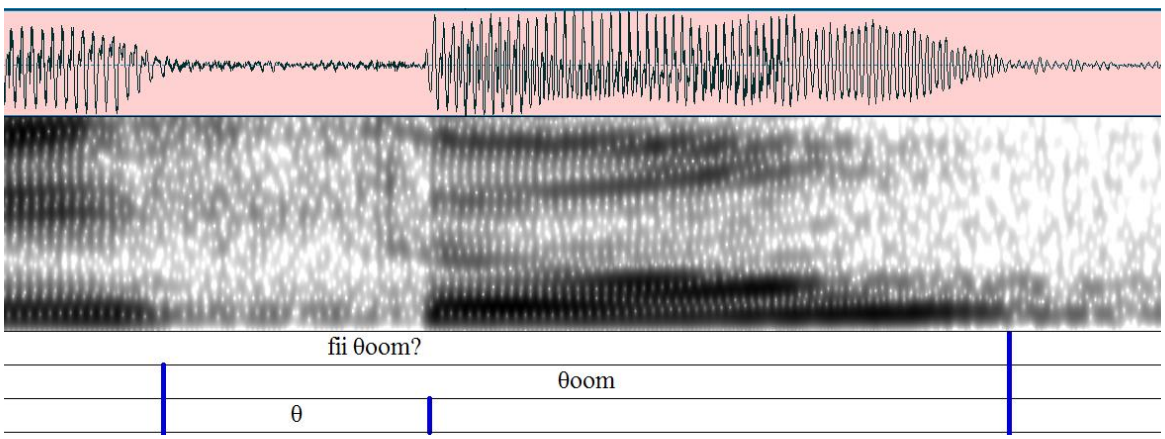


Figure 7.27 Spectrogram and waveform of the voiceless dental /θ/realized as a fricative in the word *θoːm* in FDH-directed speech

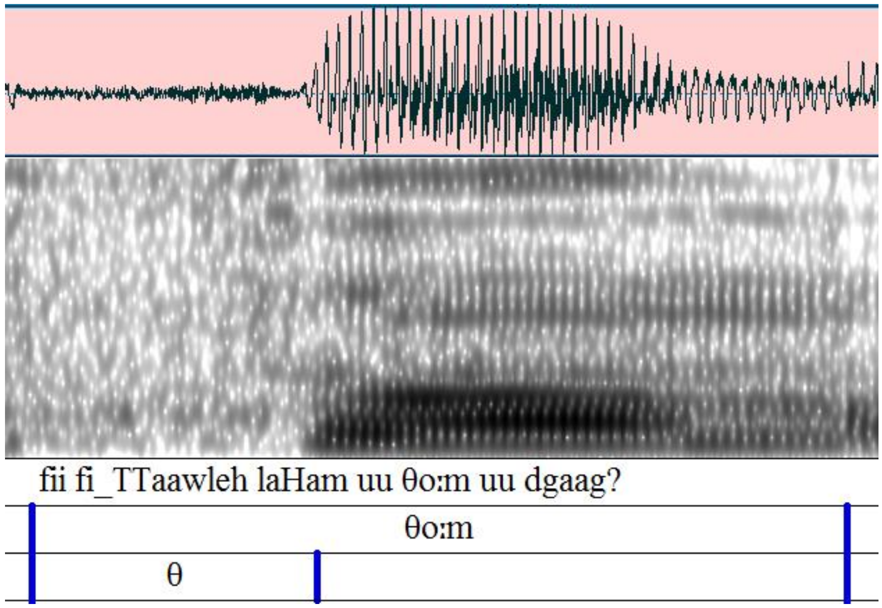


Figure 7.28 Spectrogram and waveform of the voiceless dental /θ/realized as a fricative in the word *θoːm* in ADS

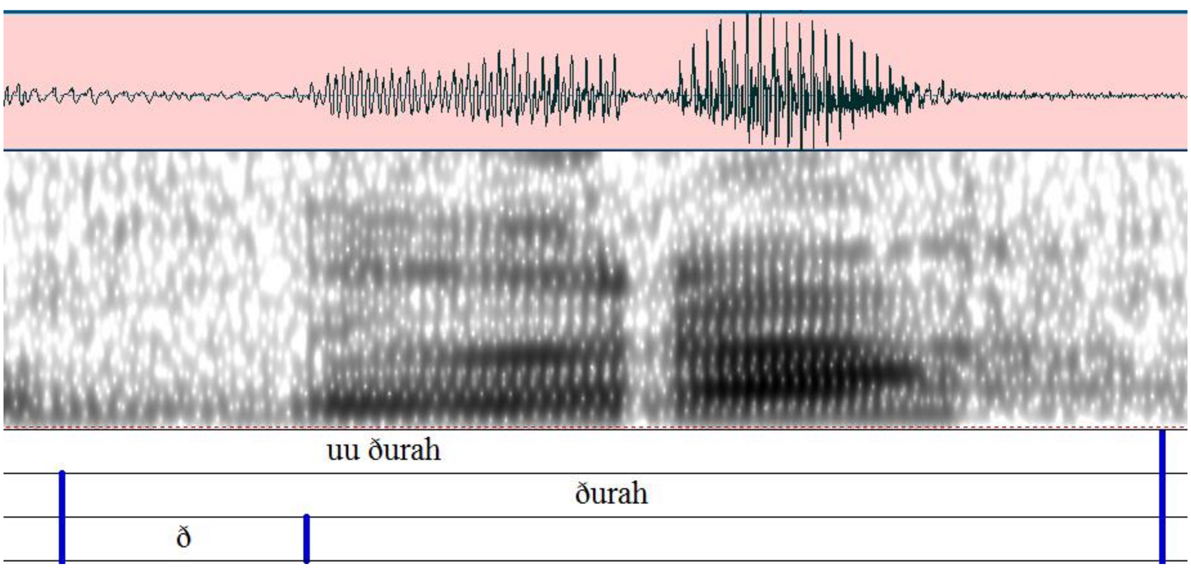


Figure 7.29 Spectrogram and waveform of the voiced dental /ð/realized as a fricative in the word *ðurah* in FDH-directed speech

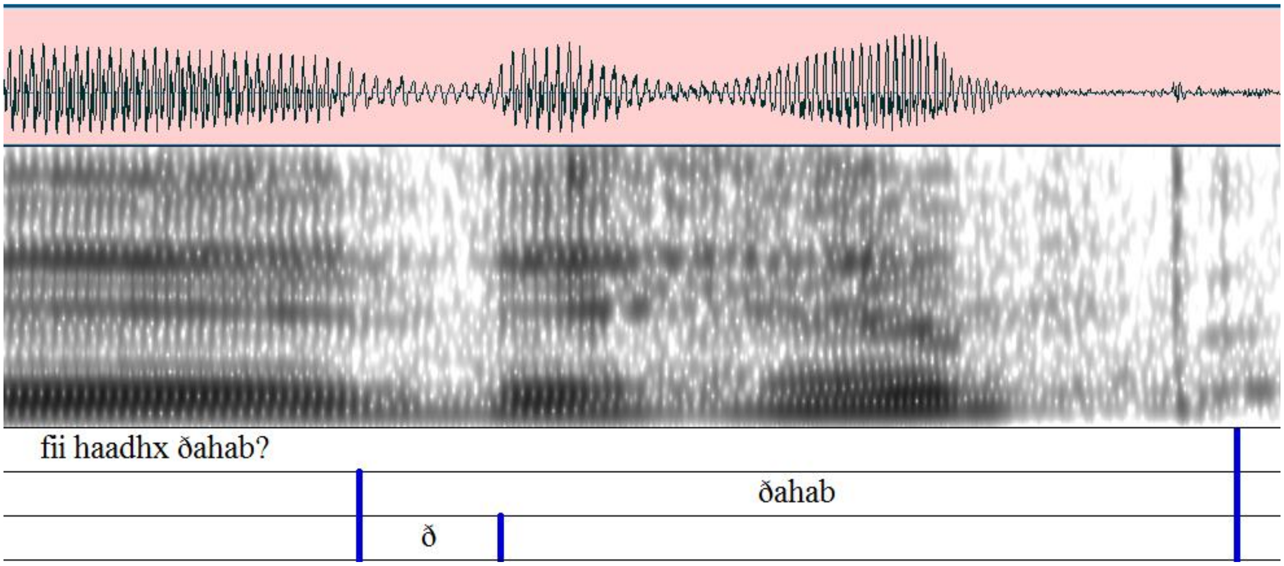


Figure 7.30 Spectrogram and waveform of the voiced dental /ð/realized as a fricative in the word *ðahab* in ADS

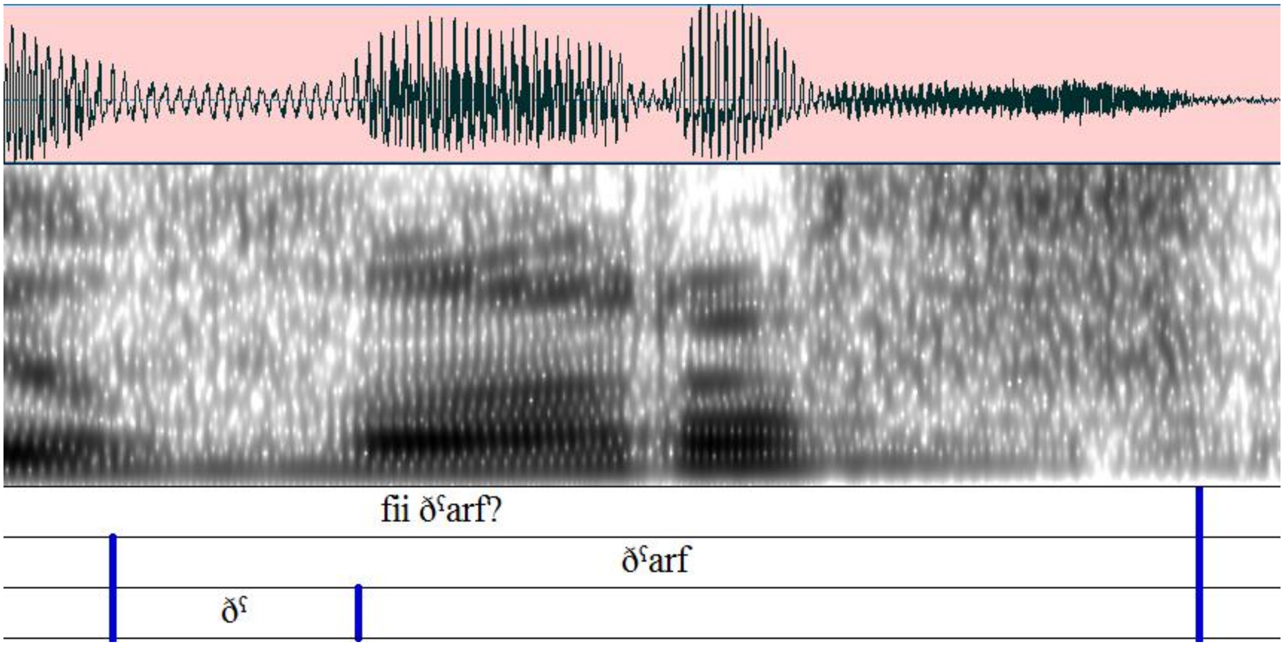


Figure 7.31 Spectrogram and waveform of the voiced dental emphatic /ðˤ/realized as an emphatic dental fricative in the word *ðˤarf* in FDH-directed speech

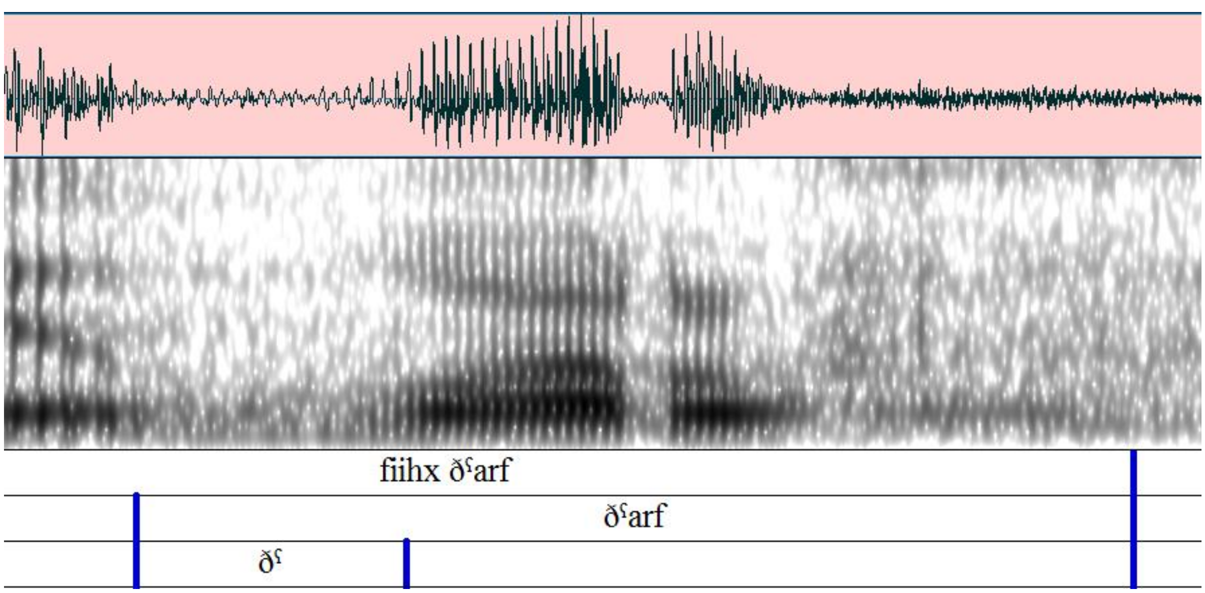


Figure 7.32 Spectrogram and waveform of the voiced dental emphatic /ðˤ/realized as an emphatic dental fricative in the word *ðˤarf* in ADS

- Stop-like

The dentals appeared in some cases produced as dental stops by the NSs. To my knowledge, there are no descriptions in the Arabic literature of dentals having stop-like realizations, but literature on other languages suggests that this is common (e.g. Lee 1992; Jurafsky *et al.* 1998; Bell *et al.* 2003; Zhao 2010). For example, In English, /ð/ has been reported to undergo modification from its full form especially when word initially and in function words. One variant that English /ð/ has been found to exhibit in casual speech is stop-like (Zhao 2010). An example of /θ/ appearing as a voiceless stop is illustrated in Figures 7.33 and 7.34. The spectrograms show random weak noise throughout the segment pointing to a mildly fricative manner of articulation and clear multiple bursts followed by a gap before the onset of the following vowel.

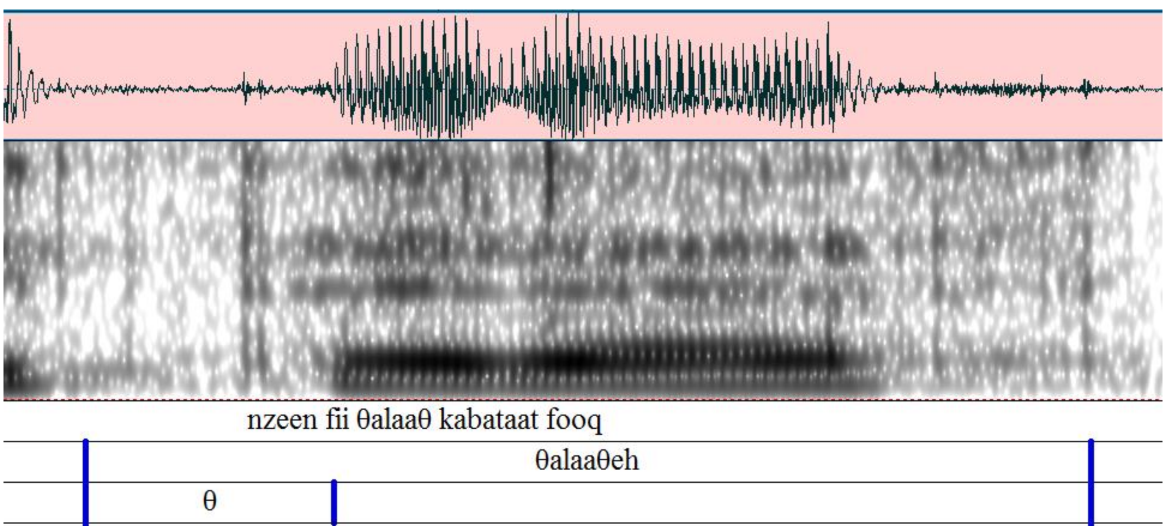


Figure 7.33 Spectrogram and waveform of the voiceless dental /θ/realized as a voiceless stop in the word *θala:θeh* in FDH-directed speech

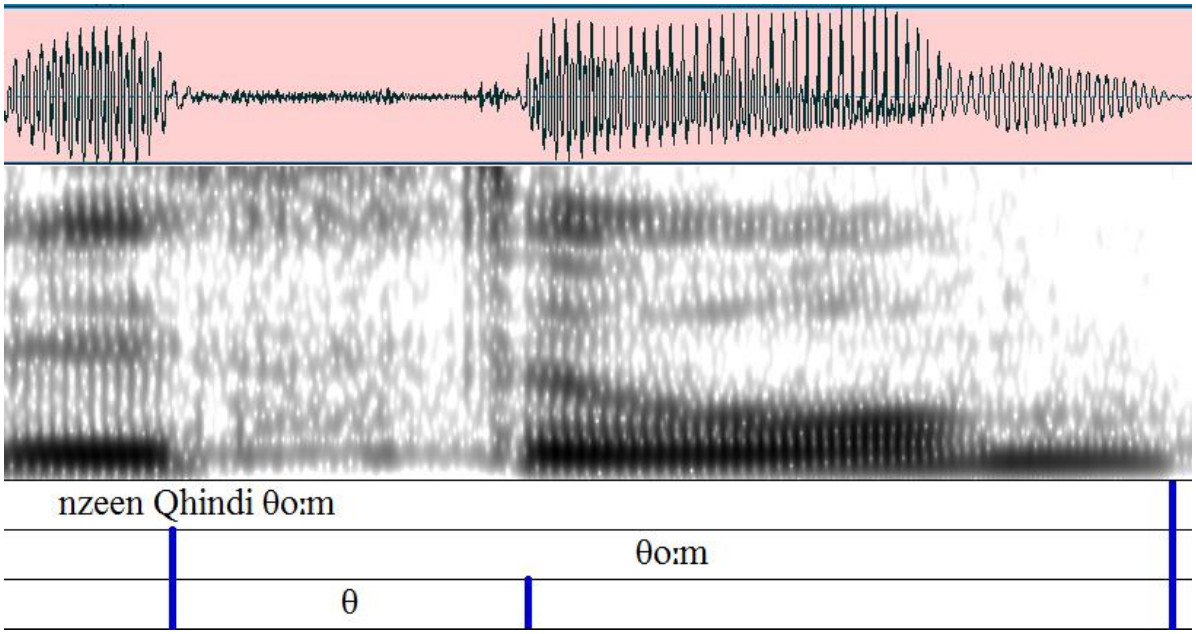


Figure 7.34 Spectrogram and waveform of the voiceless dental /θ/realized as a voiceless stop in the word *θoːm* in ADS

Figures 7.33 and 7.34 illustrate examples of /ð/ realized as a stop in both FDH-directed speech and ADS. The spectrograms show very weak noise in the lower formants indicating voicing followed by a clear burst. In figures, /ðˤ/ is articulated with an occlusion. The spectrograms show lower energy pointing to voicing followed by a burst. In the current study, all the dentals examined exhibited a stop-like realization in some of their articulations. Hence, as long as this pattern appeared in both speech styles, it seems to be a feature of the NSs dialect rather than a result of who the interlocutor was. As to what linguistic factors could trigger a stop-like realization of the dentals, the current study has not made any attempts to look into this as this was beyond the scope of this research. However, this could be an interesting area for future research on Arabic dialects to look at.

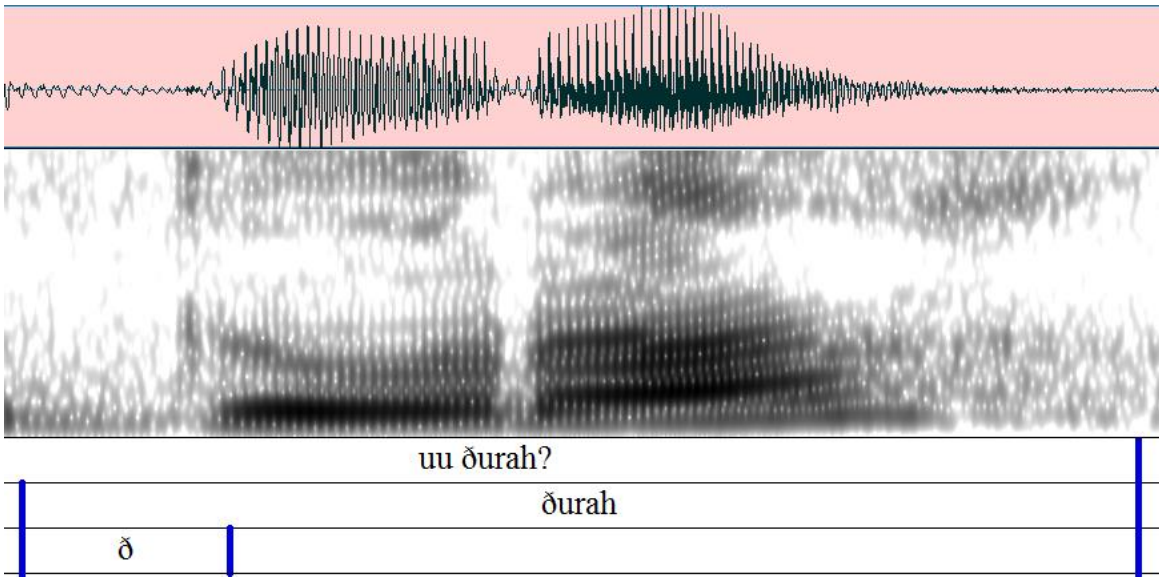


Figure 7.35 Spectrogram and waveform of the voiced dental /ð/realized as a voiced stop in the word *ðurah* in FDH-directed speech

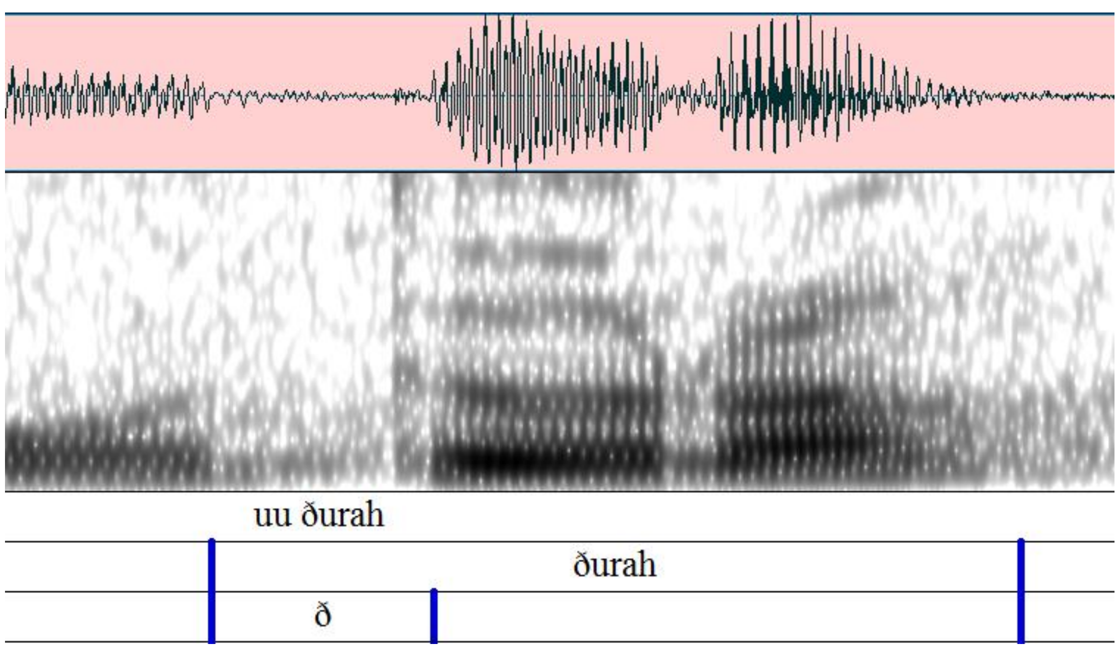


Figure 7.36 Spectrogram and waveform of the voiced dental /ð/realized as a voiced stop in the word *ðurah* in ADS

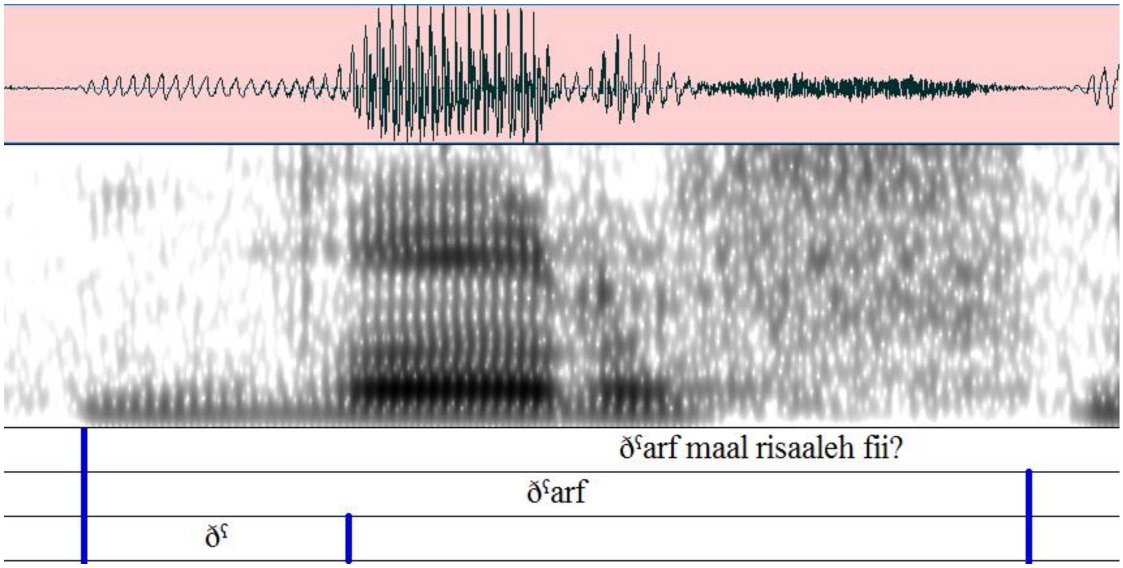


Figure 7.37 Spectrogram and waveform of the voiced dental emphatic /ðˤ/realized as a voiced emphatic dental stop in the word *ðˤarf* in FDH-directed speech

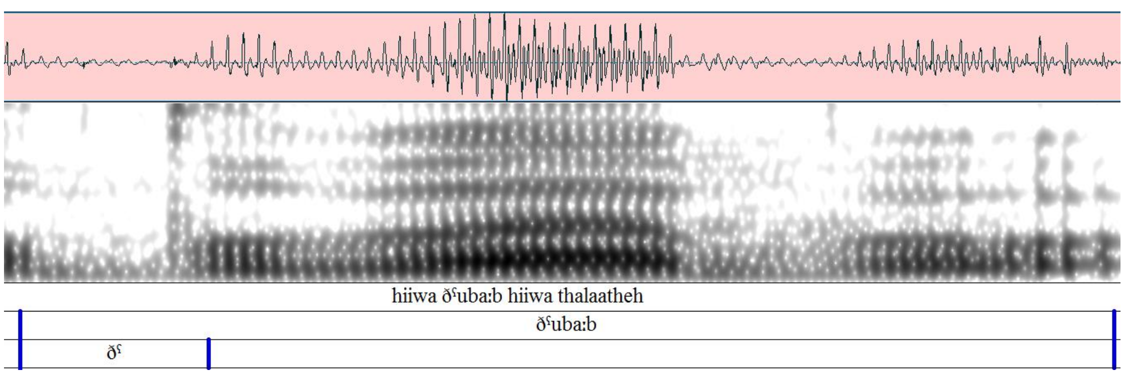


Figure 7.38 Spectrogram and waveform of the voiced dental emphatic /ðˤ/realized as a voiced emphatic dental stop in the word *ðˤba:b* in ADS

7.2.5 Summary and discussion

Auditory and spectrographic analyses of the target consonants revealed the absence of accommodation in FDH-directed speech. This finding does not provide support to H1 that states that NSs will simplify marked consonants in FDH-directed speech. H4, which states that FDH-directed speech will contain less simplified consonants as FDHs' LoR increases, was not tested due to the absence of modifications of consonants in FDH-directed speech. Auditory evidence showed that NSs produced /t ˤ, ðˤ, sˤ, χ, ʁ, q, ħ, ʕ, θ, ð / in a native-like manner, without any adaptations that would suggest convergence to the FDHs' speech despite the absence of these consonants from the FDHs' L1s. Spectrographic evidence confirmed this finding by revealing no difference in terms of acoustic patterning of the target consonants across speech style. A number of consonants showed variation in their acoustic configuration in spectrograms including the pharyngeal /ʕ/, the voiced uvular /ʁ/, and the three dental consonants /θ/, /ð/ and /ðˤ/. The pharyngeal /ʕ/ appeared in three different realizations in the speech of the NSs: approximant, fricative and stop. All three realizations have been reported to occur in some Arabic dialects (Al-Ani 1970; Butcher and Ahmad 1987; Ghazeli 1977; Holes 2004). Most importantly to the current study, the different realizations of /ʕ/ in FDH-directed speech do not point to any modification of this consonant that could show properties of a less marked pronunciation. They rather show possible and existing realizations of this variable reported elsewhere in the Arabic literature.

The voiced uvular /ʁ/ appeared in three different realizations in both FDH-directed speech and ADS: vowel-like, fricative and stop-like. The vowel-like and fricative realizations of this sound have been reported in the literature in some Arabic dialects (e.g. Al-Ani 1970; Bin-Muqbil 2006). All forms of /ʁ/ realization occurred in both speech styles, and thus do not indicate any sort of simplification of this sound. The three dental consonants /θ/, /ð/ and /ðˤ/ appeared in two different realizations in both FDH-directed speech and ADS: fricative and stop-like. The fricative manner of articulation of these consonants is the one that has been generally reported in the literature (Al-Ani 1970). The three dental consonants appeared more often as a fricative than as a stop in both FDH-directed speech and ADS. Most importantly, the stop realization of the dental fricatives does not point to any phonetic modification intended to simplify these consonants to FDHs. Further evidence for the absence of any simplification in FDH-directed speech can be drawn from findings on the acoustic analysis of the emphatic/plain stops and fricatives. Results on spectral moments of the emphatic/plain consonants and the formant frequencies of the vowels following them showed similar trends in both FDH-directed speech and ADS as will be further discussed in the following section. Further discussion of these results will be provided in Chapter 9.

7.3 Realization of Consonant Clusters in FDH-directed Speech

7.3.1 Onset clusters

Table 7.5 presents proportions of deletion in FDH-directed speech and ADS. The total number of tokens in FDH-directed speech is roughly double the number in ADS. From the descriptive statistics, it is noticeable that NSs in this study prefer to retain the short unstressed vowel in onset consonant clusters rather than deleting it in both ADS and FDH-directed speech (60% and 73.14% respectively). In other words, there was a general trend among NSs to diverge from syncope regardless of the interlocutor, as the percentage of vowel deletion is less than 50% in both speech styles.

Consonant Cluster (CC)	FDS				ADS		
	Keyword	Tokens	Deletion		Tokens	Deletion	
			%Yes	%No		%Yes	%No
dg-	/dgaːg/	37	0	10.57	31	1.11	16.11
ðˤb-	/ðˤbaːb/	23	0.86	5.71	6	1.67	1.66
gd-	/gdaːr/	61	6.86	10.57	30	7.22	9.44
kt-	/ktaːb/	82	5.14	18.28	31	5.55	11.66
lħ -	/lħam/	25	0	7.14	25	0	13.88
sˤʁ -	/sˤʁiːr/	55	4	11.71	10	3.89	1.67
st-	/staːrah/	67	10	9.14	47	20.55	5.55
Total		350	26.86%	73.14%	180	40%	60%

Table 7.5 Descriptive statistics for the deletion proportions of onset CCs across speech style

GLMM demonstrated that vowel deletion was significantly lower in FDH-directed speech compared to ADS as Table 7.6 and Figure 7.39 show (p<0.01). This indicates that in FDS-directed speech, the short vowel was retained more often than in ADS, pointing to a tendency by NSs to simplify onset consonant clusters more often when addressing FDHs.

Predictor	Estimate	Std.Error	Z value	Pr(>|z|)
ADS (Baseline)	1.74	0.98	1.76	
FDS	1.23	0.9	4.35	**1.35e-05 *****
Random Effects	variance	Std.Deviation		
Speaker	2.22	1.49		
Target Word	5.31	2.305		

Table 7.6 GLMM results for the effect of speech style on the deletion of onset CCs

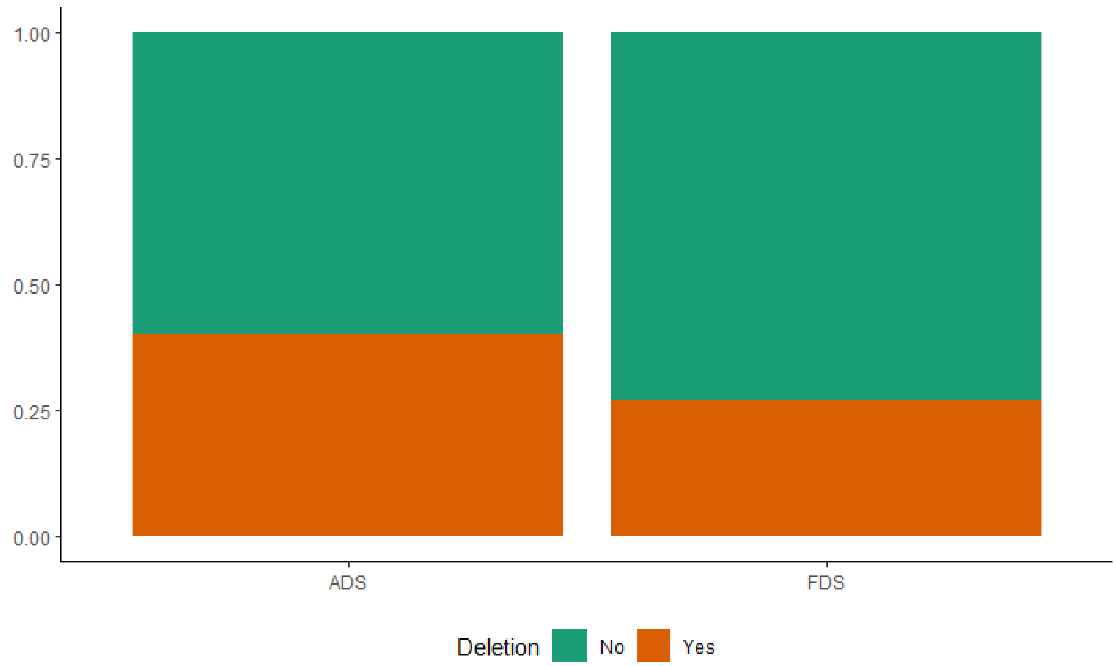


Figure 7.39 The effect of speech type on vowel deletion in onsets

To rule out any effects of condition order in the significant results above, Table 7.7 shows that vowel deletion was higher in FDH-directed speech in both condition orders. GLMM demonstrated that in condition order 1, in which NSs interacted with FDHs first in the task, vowel deletion was significantly lower in FDH-directed speech compared to ADS as shown in Figure 6.40 and table 6.8 ($p<0.01$). On the contrary, in condition order 2, in which FDH-directed speech was produced second, FDH-directed speech and ADS were statistically comparable with regard to vowel deletion ($p>0.05$).

Predictor	Estimate	Std.Error	Z value	Pr(>|z|)	%Yes
Condition Order 1 (FDS/ADS)					
ADS (Baseline)	0.77	0.95	0.81		44.44
FDS	2.303	0.48	4.76	**1.88e-06 *****	20.66
Random Effects	variance	Std.Deviation			
Speaker		1.17	1.08		
Target Word		4.75	2.18		
Predictor	Estimate	Std.Error	Z value	Pr(>|z|)	%Yes
Condition Order 2 (ADS/FDS)					
ADS (Baseline)	2.65	1.42	1.86		35.55
FDS	0.06	0.41	0.16	0.87	31.5

Random Effects	variance	Std.Deviation
Speaker	5.05	2.24
Target Word	8.08	2.84

Table 7.7 GLMM results for the effect of condition order on modifications of onset CCs

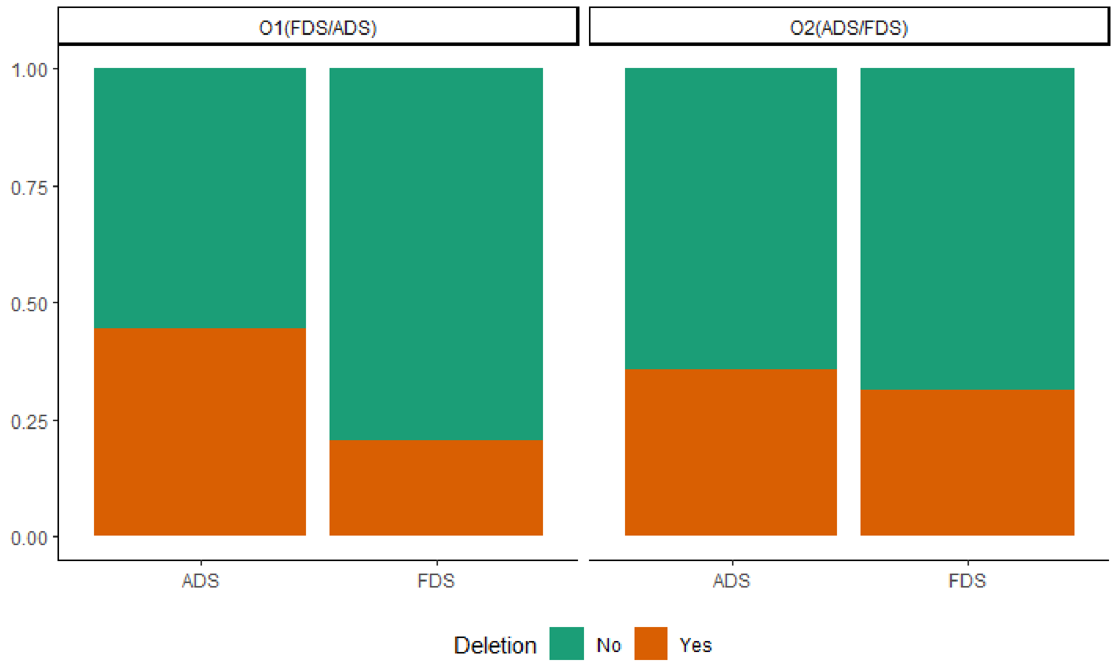


Figure 7.40 The effect of condition order on modifications of onset CCs

7.3.1.1 The role of LoR

GLMM revealed that LoR did not play a significant role in how frequently vowel deletion occured in FDH-directed speech as shown in Table 7.8.

	Estimate	Std.Error	Z value	Pr(>|z|)
(Intercept)	3.201	1.12	2.83	
LoR	0.101	0.091	1.11	0.26
Random Effects	variance	Std.Deviation		
Speaker	4.26	2.06		
Target Word	5.12	2.26		

Table 7.8 GLMM results on the effect of LoR on vowel deletion in FDH-directed speech

7.3.2 Coda clusters

Table 7.9 and Figure 7.41 present proportions and tokens of epenthesis in FDH-directed speech and ADS for each target word. They illustrate that vowel epenthesis in coda consonant clusters was very rare in both FDH-directed speech and ADS (3.72% and 1.03%, respectively). Cluster simplification was attested in one target word only, that is /χubz/ 'bread'. It is worth noting that where epenthesis occurred it was in a cluster that violated the sonority sequencing principle (SSP). According to the SSP, more sonorous segments stand closer to the syllable peak compared to less sonorous segments. The cluster /-bz/ consists of a stop+fricative. Since stops are assumed to be less sonorous than fricatives based on Universalist theories, the previous cluster is ill-formed based on sonority scale and thus violates the SSP (Daland *et al.* 2011). On the other hand, all the other clusters are well-formed as they consist of segments clustered in a way that adheres to the SSP (most sonorous to least sonorous). This could be one reason why epenthesis occurred in the /-bz/ cluster and not in the others, that is to conform to the SSP.

Cluster	FDS				ADS		
	Keyword	Tokens	Epenthesis		Tokens	Epenthesis	
			%Yes	%No		%Yes	%No
-bz	/χubz/	31	3.72	15.53	20	1.03	19.59
-lb	/kalb/	44	0	27.32	28	0	28.86
-nt	/bɪnt/	41	0	25.46	29	0	29.89
-rf	/ðˤarf/	45	0	27.95	20	0	20.62
Total		201	3.72%	96.27%	97	1.03%	98.97%

Table 7.9 Descriptive statistics for epenthesis in coda CCs across speech style

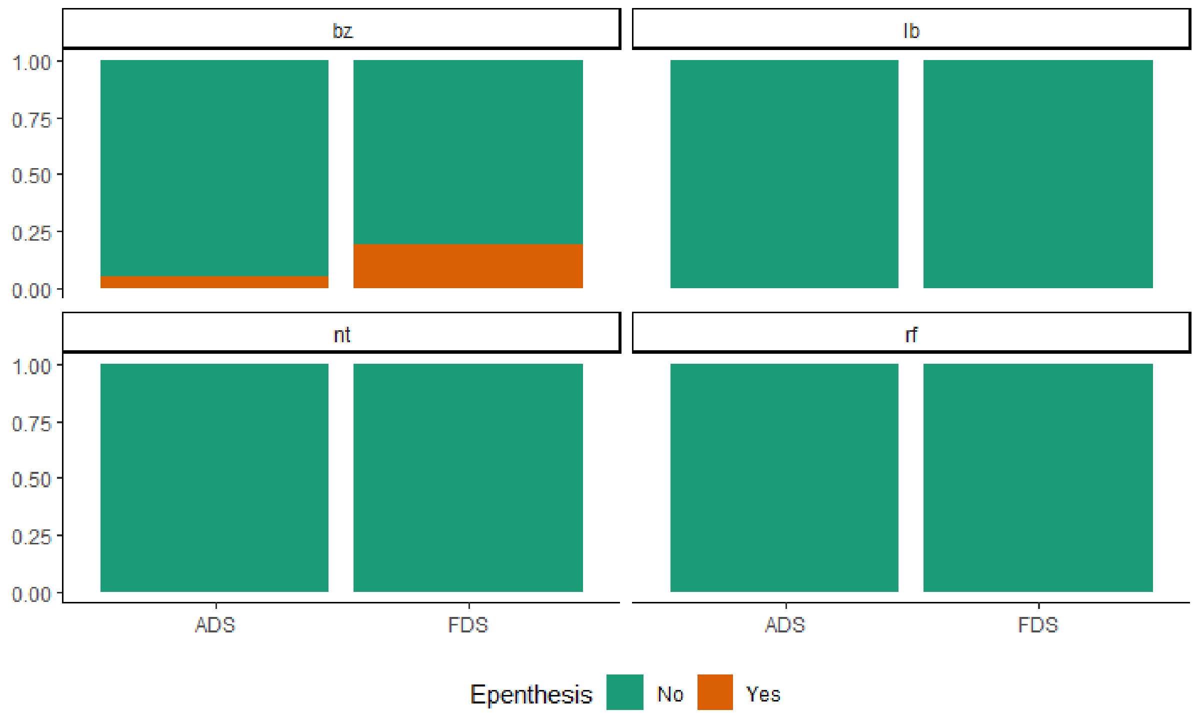


Figure 7.41 Proportions of epenthesis across speech style in each target CC

GLMM demonstrates that the difference between ADS and FDH-directed speech in the simplification of coda consonant clusters was significant as evident in Table 6.10 and Figure 7.42 ($p<0.05$). This result might not be meaningful as it was based on one target word only and very few tokens. Hence, we cannot assume that FDH-directed speech is undergoing consonant cluster modification. Since the frequency of epenthesis was very small and might not be meaningful, the effect of LoR and condition order were not examined.

	Estimate	Std.Error	Z value	Pr(>|z|)	%Yes
ADS (Baseline)	58.36	20.87	2.79		1.03
FDS	-14.46	6.61	-2.18	**0.02***	3.72
Random Effects	variance	Std.Deviation			
Speaker	2650	51.48			
Target Word	1014	31.84			

Table 7.10 GLMM results for the effect of speech style on coda CC epenthesis

Figure 7.42 The effect of speech style on vowel epenthesis in coda CCs

7.3.3 Summary and Discussion

H2 which states that FDH-directed speech will involve less marked onset and coda consonant clusters was partially supported. The complex syllable structure in FDH-directed speech was found to be modified more often compared to that in ADS. Syncope (deletion of the short vowel in an unstressed syllable) was found to be less evident in FDH-directed speech than ADS, indicating a tendency of simplification of the consonant clusters in speech directed to FDHs. Despite significant differences between FDH-directed speech and ADS with regard to vowel deletion in onset consonant clusters, there was a general trend of short vowel deletion of the unstressed syllables of the target words attested among NS participants. This can be interpreted in relation to findings from Ambu Saidi (2019) who examines whether Nizwa migrants adhere to the rule of syncope due to contact with supralocal variants in the Capital Muscat. Similar to findings from this study, she found that there was a general trend of vowel deletion among all her participants that had an overall percentage of less than 50%. Ambu Saidi explained this pattern as an indication of participants' change of social status and avoidance of stigmatization; emphasizing their educated and modern status and avoiding being perceived as rural. Hence, findings from this study could be explained in similar veins to those of Ambu Saidi when it comes to explaining the previous trend in ADS. NSs might have wished to emphasize their educated urban status when interacting with the NS adult, especially that she was a stranger to them. However, NSs' production pattern when addressing FDHs cannot be explained in the same way. This is due to the different social status the FDH had compared to that of the NS adult. The tendency towards avoiding syncope in FDH-directed speech can be attributed to FDHs limited Arabic capacity and perceived foreign-ness. Further discussion of this will be provided in Chapter 8.

Findings from the previous section did not provide support to H4 which states that FDH-directed speech will contain less simplified consonants and consonant clusters or less hyperarticulation as FDHs' LoR increases. LoR did not predict any modification of the syllable in FDH-directed speech. Hence, the NS modified her onset consonant clusters when interacting with her FDH regardless of how experienced the FDH was in Arabic. Modifications of coda consonant clusters were less prevalent than those of onset consonant clusters and quite rare in both FDH-directed speech and ADS.

7.4 Acoustic Examination of Consonants in FDH-directed speech

This section reports on results concerning the acoustic analyses of the plain/emphatic fricatives /s/ and /sˤ/ as well as the plain/emphatic stops /t/ and /tˤ/. This examination was set to test the following hypotheses:

H3. NSs will hyperarticulate their speech acoustically and prosodically to enhance its clarity for their FDH interlocutors.

H3a. NSs will enhance the contrast between plain/emphatic consonants by hyperarticulating the prominent acoustic correlates that the NSs in the current study use to distinguish emphatic from plain consonants.

H3b. Consonant intensity will be higher in FDH-directed speech.

H3c. Fricative duration will be longer and VOT of stop consonants will be longer in FDH-directed speech.

H4. Simplification and hyperarticulation in FDH-directed speech will vary based on FDHs' foreign accentedness score, LoR and religion.

To test H3 and its sub-hypotheses, the study examined spectral moments (center of gravity, standard deviation, skewness, and kurtosis), duration and intensity of the plain/emphatic consonants as well as the three formant frequencies (F1, F2, F3) of the following vowels at the onset position.

7.4.1 Data cleaning

Since the task used to elicit NSs' speech lent itself to spontaneous interactions prompted by objects in the pictures, the target plain/emphatic consonants were subject to phonological processes from the neighboring environment. One major phonological process that affected the target consonants was that of assimilation of a preceding definite article, leading the consonants to geminate. Traditional Arab grammarians have divided Arabic consonants into two categories: (1) Fourteen non-coronal consonants called *al-ḥurūf al-qamarīya* ('letters of the moon') including /b ʤ k q ʔ f χ ʁ ħ ʕ h m w j/, (2) Fourteen coronal consonants known as *al-ḥurūf al-šamsīya* ('letters of the sun') including /t tˤ d dˤ θ ð ðˤ s sˤ z ʃ l n r/ (Kenstowicz 1994). Most accounts on Arabic dialects report that when coronal consonants in nominal categories are preceded by the definite article, /ʔal/ (/ʔɪl/ in Nizwa Arabic), they are produced as geminates due to a complete assimilation of the lateral /l/ to the following coronal consonant in manner, aspect or place of articulation (e.g. Salib 1981; Kenstowicz 1994; Youssef 2013; Watson 2002; Kambuziya 2007). Since the target plain/emphatic consonants in this study are coronals, they were subject to gemination when preceded by a definite article as in (7.1):

(7.1) ʔal + taːg ⟶ ʔat-taːg 'the crown'

Table 7.11 illustrates the number of tokens containing the plain/emphatic consonants in FDH-directed speech and ADS based on how the target plain/emphatic sounds were realized in the speech of the NSs. The geminated sounds are the result of an assimilation process of a preceding definite article.

Target Word	Target Sound realization	FDH-directed speech	ADS	Total
/taːg/	[t] [tt]	36 1	24 0	61
/talafuːn/	[t] [tt]	31 5	19 1	56
/tˤaːwleh/	[tˤ] [tˤtˤ]	40 35	11 24	110
/tˤamaːtˤ/	[tˤ] [tˤtˤ]	21 1	23 4	49
/samak/	[s] [ss]	38 3	25 1	67
/sayaarah/	[s] [ss]	33 3	23 3	62
/sˤaħan/	[sˤ] [sˤsˤ] [s]	38 4 4	21 0 6	73
/ sˤaabuːn/	[sˤ] [sˤsˤ]	30 0	23 4	57
Total		323(60.4%)	210(39.2%)	535

Table 7.11 Number of tokens of the plain/emphatic target sounds in terms of how NSs realized the target consonants in FDH-directed speech and ADS

In addition to the aforementioned variation in the production of the target consonants, vowels following some of the target consonants were also subject to variation between and within speakers. Table 7.12 shows the number of tokens containing the plain/emphatic consonants in FDH-directed speech and ADS based on how the following vowel was realized by NSs. It demonstrates that vowels following the plain alveolars /t/ and /s/ in the words /talafuːn/ and /sajjaːrah/ varied in the NSs' speech between [a] and [i] as the examples in 7.2 below further explain:

(7.2)

(a) /talafuːn/ → [**ta**lafuːn] or [**ti**lifoːn]

(b) /sajjaːrah/ → [**sa**jjaːrah] or [**si**jjaːrah]

Target Word	Following vowel realization	FDH-directed speech	ADS	Total
/taːg/	[aː]	37	24	61
/talafuːn/	[a] [i]	13 23	7 13	56
/ tˤaːwleh/	[aː]	75	35	110
/tˤamaːtˤ/	[a]	22	27	49
/samak/	[a]	41	26	67
/sajjaːrah/	[a] [i]	27 9	19 7	62
/sˤaħan/	[a]	46	27	73
/ sˤaːbuːn/	[aː]	30	27	57
Total		323 (60.4%)	210 (39.2%)	535

Table 7.12 Number of tokens of the plain/emphatic target sounds in terms of how NSs realized the following vowel in FDH-directed speech and ADS

A well-known phenomenon in speech production is that the vocal tract shape of the consonant is influenced by the adjacent vowel or consonant (Stevens 1994). Due to this effect, modifications to the formant frequencies in the nearest vicinity of the consonant are expected to occur (ibid). Ultimately, to have comparable samples of the target consonants and their following vowel in the statistical analysis of the main study, tokens produced with the high front vowel [i] following the target consonants have been removed from the analysis. Moreover, all consonants that were geminated due to assimilation of a preceding definite article have been removed from the analysis. Tokens of the word /sˤaħan/ that were realized as /saħan/, with the plain /s/, were also excluded from the analysis. Table 7.13 shows the ultimate number of tokens included in statistical analyses in both FDH-directed speech and ADS.

Target Consonant	FDH-directed speech	ADS
/s/	62	42
/sˤ/	68	44
/tˤ/	60	34
/t/	48	31
Total	238 (61.2%)	151(38.8%)

Table 7.13 Number of tokens for each target consonant included in statistical analysis

The following section will provide results on the acoustic properties of /s/ and /sˤ/.

7.4.2 Realization of plain/emphatic fricative consonants

Linear Mixed Effect Models (LMEMs) demonstrated that there was no main effect of sound type on any of the four spectral moments, indicating that /s/ and /sˤ/ were not distinguished based on spectral moments in the speech of the NSs in the current study as Table 7.14 shows ($p>0.05$). Furthermore, results showed that the plain/emphatic fricatives were not distinct with regard to duration or intensity ($p>0.05$). From this, it is safe to say that spectral moments, duration and intensity are not reliable acoustic correlates of fricative emphasis in the speech of the NSs in the current study. Thus, we cannot rely on these metrics to predict the enhancement of the /s/ vs. /sˤ/ contrast in FDH-directed speech. NSs in the current study may maximize the distance between /s/ and /sˤ/ by changing F1 and F2 of vowels following these continuants (Section 7.4.2.6).

With regard to the effect of speech style (ADS vs. FDH-directed speech), skewness of /sˤ/ was significantly higher in FDH-directed speech compared to the baseline (speech style: ADS, sound type: /sˤ/) ($p<0.05$). Also, the duration of /sˤ/ was significantly shorter in FDH-directed speech compared to the baseline ($p<0.01$). The effect of speech style on of all the other metrics was non-significant. Interaction between sound type and speech style was also non-significant in all metrics examined. Detailed results of *post hoc* pairwise comparison analyses for each metric will be provided in the following sections.

Metric	Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
Centre of Gravity (Hz)	Intercept (ADS, /sˤ/)	7961.69	290.31	45.75	27.42	
	Sound: /s/	74.77	268.18	196.40	0.27	0.78
	Speech: FDS	-90.79	239.89	195.75	-0.37	0.705
	/s/: FDS	-209.75	347.48	196.65	-0.604	0.54
Random effects	speaker	Variance 1071905	Std.Deviation 1035			
Std. Deviation (Hz)	(Intercept)	2848.4	159.02	63.43	17.91	
	Sound: /s/	-117.14	168.75	198.89	-0.69	0.48
	Speech: FDS	103.66	151.01	198.01	0.68	0.49
	/s/: FDS	204.06	218.60	199.32	0.93	0.35
Random effects	speaker	Variance 247318	Std.Deviation 497.3			
Skewness (Hz)	(Intercept)	-0.87	0.13	62.08	-6.34	
	Sound: /s/	-0.15	0.14	198.43	-1.05	0.29
	Speech: FDS	0.29	0.13	197.54	2.24	**0.02***
	/s/: FDS	0.14	0.19	198.87	0.74	0.45
Random effects	speaker	Variance 0.18	Std.Deviation 0.18			
Kurtosis (Hz)	(Intercept)	1.99	0.43	71.03	4.59	
	Sound: /s/	0.03	0.48	199.13	0.069	0.94
	Speech FDS	-0.15	0.43	198.09	-0.36	0.71
	/s/: FDS	-0.16	0.63	199.73	-0.26	0.79
Random effects	speaker	Variance 1.57	Std.Deviation 1.25			
Duration (ms)	(Intercept)	114.44	3.72	57.64	30.71	
	Sound: /s/	-2.508	3.76	198.54	-0.66	0.506
	Speech FDS	-11.48	3.37	197.77	-3.404	**0.00 *****
	/s/: FDS	6.29	4.88	198.88	1.28	0.19
Random effects	speaker	Variance 151.0	Std.Deviation 12.29			
Intensity (dB)	(Intercept)	51.88	0.88	49.09	58.62	
	Sound: /s/	0.808	0.85	196.58	0.94	0.34
	Speech FDS	0.87	0.76	195.84	1.14	0.25
	/s/: FDS	-0.56	1.107	196.88	-0.51	0.609
Random effects	speaker	Variance 9.30	Std.Deviation 3.05			

Table 7.14 Results of LMEM models for the effect of speech style and sound type on the realization of the plain/emphatic fricatives

7.4.2.1 Spectral moments

- **Center of Gravity (CoG):** CoG of both /s/ and /sˤ/ was lower in FDH-directed than in ADS, though this was insignificant as shown in Figure 7.43 and Table 7.15 (see Emmean). Lower CoG suggests a concentration of energy in lower frequencies (Maniwa *et al*. 2009). *Post hoc* tests revealed that the pattern of ADS and FDH-directed speech was not entirely parallel as shown in Table 7.15. ADS showed a pattern in which CoG decreases as place moves back (CoG for /sˤ/ is 74.8 Hz lower than that for /s/). However, FDH-directed speech was different in that its pharyngealized fricative had a higher CoG than its plain alveolar fricative (135 Hz higher). However, a sound × speech style interaction did not prove significant based on the LMEM, indicating that the previous patterns were not robust.
- **Standard Deviation (SD):** Contrary to CoG, SD of both /s/ and /sˤ/ was higher for FDH-directed speech than for ADS as Figure 7.43 and Table 7.15 show (see Emmean). Higher SD indicates a concentration of energy towards higher frequencies. Furthermore, *post hoc* tests revealed that SD of /s/ was lower in ADS compared to FDH-directed speech, but this was insignificant ($p=0.05$).
- **Skewness:** Table 7.15 and Figure 7.43 demonstrate that skewness of both /s/ and /sˤ/ was significantly higher in FDH-directed speech than in ADS based on *post hoc* tests ($p<0.05$). This indicates that spectra of FDH-directed speech had a more negative tilt and a concentration of energy towards lower frequencies (Jongman *et al.* 2000).
- **Kurtosis:** Kurtosis of /s/ and /sˤ/ is lower in FDH-directed speech compared to ADS as shown in Figure 7.43 and Table 7.15 (see Emmean). But, the difference was insignificant. Lower kurtosis indicates a flat spectrum with less defined peaks (Jongman *et al.* 2000). The pattern of kurtosis in FDH-directed speech and ADS was not completely analogous. In ADS, kurtosis of /sˤ/ was 0.03 Hz lower than that for /s/, whereas in FDH-directed speech, kurtosis for /sˤ/ was 0.13 higher than that for /s/. These differences, however, proved to be negligible based on the sound ×speech style interaction revealed by the LMEM.

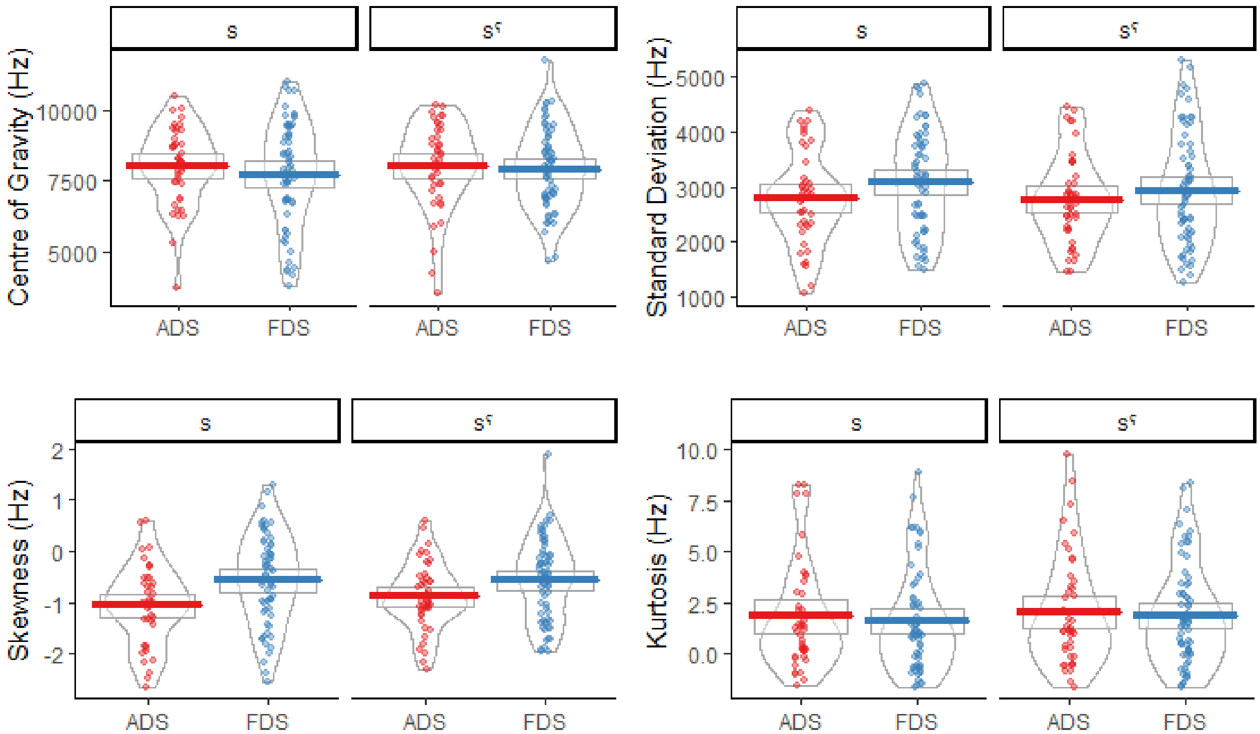


Figure 7.43 Spectral Moments of /s/ and /sˤ/ in FDH-directed speech and ADS

Metric							
	Sound	Register	Emmean	SE	df	lower.CL	upper.CL
	sˤ	ADS	7962	296	49.5	7367	8556
		FDS	7871	276	37.9	7312	8430
	s	ADS	8036	297	50.4	7440	8633
		FDS	7736	280	39.7	7171	8301
Center of Gravity	Level	Contrasts	estimate	SE	df	t.ratio	p.value
	sˤ	ADS-FDS	90.8	242	199	0.37	0.707
	s	ADS-FDS	300.5	252	201	1.19	0.23
	ADS	sˤ -s	-74.8	270	200	-0.27	0.78
	FDS	sˤ -s	135.0	224	202	0.603	0.54
	sˤ	ADS	2848	162	66.0	2525	3172
		FDS	2952	147	46.6	2656	3248
	s	ADS	2731	163	68.1	2407	3056
		FDS	3039	150	49.7	2738	3340
Standard Deviation	Level	Contrasts	estimate	SE	df	t.ratio	p.value
	sˤ	ADS-FDS	-104	152	201	-0.68	0.49
	s	ADS-FDS	-308	159	203	-1.93	**0.05**
	ADS	sˤ -s	117.1	170	202	0.68	0.49
	FDS	sˤ -s	-86.9	141	204	-0.61	0.53
	sˤ	ADS	-0.88	0.14	66.0	-1.16	-0.59
		FDS	-0.58	0.12	46.6	-0.84	-0.32
	s	ADS	-1.03	0.14	68.1	-1.31	-0.75
		FDS	-0.59	0.13	49.7	-0.86	-0.33
Skewness	Level	Contrasts	estimate	SE	df	t.ratio	p.value
	sˤ	ADS-FDS	-0.29	0.13	201	-2.22	**0.02**
	s	ADS-FDS	-0.43	0.13	203	-3.16	**0.001**
	ADS	sˤ -s	0.15	0.14	202	1.04	0.29
	FDS	sˤ -s	0.01	0.12	204	0.107	0.91
	sˤ	ADS	1.99	0.44	76.7	1.11	2.87
		FDS	1.83	0.39	52.3	1.03	2.63
	s	ADS	2.03	0.44	79.8	1.14	2.91
		FDS	1.70	0.405	56.1	0.89	2.51
Kurtosis	Level	Contrasts	estimate	SE	df	t.ratio	p.value
	sˤ	ADS-FDS	0.15	0.44	202	0.36	0.71
	s	ADS-FDS	0.32	0.45	204	0.706	0.48
	ADS	sˤ -s	-0.03	0.49	203	-0.06	0.94
	FDS	sˤ -s	0.13	0.406	206	0.32	0.74

Table 7.15 Results of *post hoc* pairwise comparison tests for spectral moments of /s/ and /sˤ/

7.4.2.1.1 The effect of FDHs' foreign accented rating, LoR, and religion on spectral moments

LMEMs revealed that FDHs' foreign accentedness rating, LoR and religion did not have any significant effects on center of gravity, standard deviation and skewness ($p>0.05$). However, the correlation between the emphatic sound and foreign accentedness score significantly predicted the change in center of gravity and skewness ($p<0.05$). Also, religion was a significant predictor of fricative kurtosis ($p<0.05$).

Metric	Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
Cog	Intercept (/s/, Muslim)	8464.07	801.86	30.76	10.55	
	Target Sound:sˤ	-1042.14	536.54	110.39	-1.94	0.05
	Accent rating	350.42	228.11	32.32	1.53	0.13
	LoR	-38.48	43.09	19.93	-0.89	0.38
	Religion: Non-Muslim	-983.66	599.98	18.48	-1.63	0.11
	Target Sound*Accent rating	-504.29	200.62	111.60	-2.51	**0.01 ***
Mixed effects	speaker	Var 727474	SD 852.9			
SD	Intercept	2507.27	474.178	33.39	5.28	
	Target Sound:sˤ	-69.8	349.46	111.7	-0.2	0.84
	Accent rating	-229.25	135.27	35.55	-1.69	0.09
	LoR	2.448	24.79	19.59	0.09	0.92
	Religion: Non-Muslim	-18.09	343.23	17.57	-0.05	0.95
	Target Sound*Accent rating	11.58	130.56	113.21	0.089	0.92
Mixed effects	speaker	Var 208982	SD 457.1			
Skew	Intercept	-0.8	0.4	36.56	-2.005	
	Target Sound: sˤ	0.52	0.29	113.58	1.76	0.079
	Accent rating	-0.09	0.11	38.81	-0.79	0.43
	LoR	0.01	0.02	21.88	0.82	0.41
	Religion: Non-Muslim	0.58	0.29	19.7	1.99	0.06
	Target Sound*Accent rating	0.22	0.11	114.93	2.06	**0.04***
Mixed effects	speaker	Var 0.15	SD 0.39			
Kurt	Intercept	3.69	1.01	28.66	3.65	
	Target Sound: sˤ	-0.08	0.91	109.78	-0.08	0.92
	Accent rating	0.9	0.29	32.43	3.09	**0.004 ****
	LoR	-0.01	0.04	11.73	-0.29	0.77
	Religion: Non-Muslim	0.001	0.65	8.92	0.002	0.99
	Target Sound*Accent rating	-0.11	0.34	113.22	-0.32	0.74
Mixed effects	speaker	Var 0.37	SD 0.61			

Table 7.16 LMEMs results for the effect of FDHs' foreign accentedness, LoR and religion on spectral moment

Center of gravity

FDHs’ foreign accent score, LoR or religion did not have significant effects on fricative CoG ($p>0.05$). Figure 7.44 illustrates the significant interaction effect between fricative sound and foreign accent rating revealed by the LMEM. It shows that the more Native-like a FDH is, center of gravity of /s/ in FDH-directed speech increases. On the other hand, the more Native-like a FDH is, center of gravity of /sˤ/ in FDH-directed speech drops. Thus, the relationship between center of gravity and the type of fricative sound produced in FDH-directed speech essentially depends on FDHs’ foreign accentedness. This suggests that NSs realize a more fronted /s/ when they address FDHs’ who sound more Native-like. However, NSs’ /sˤ/ is more retracted when they address the same group of FDHs.

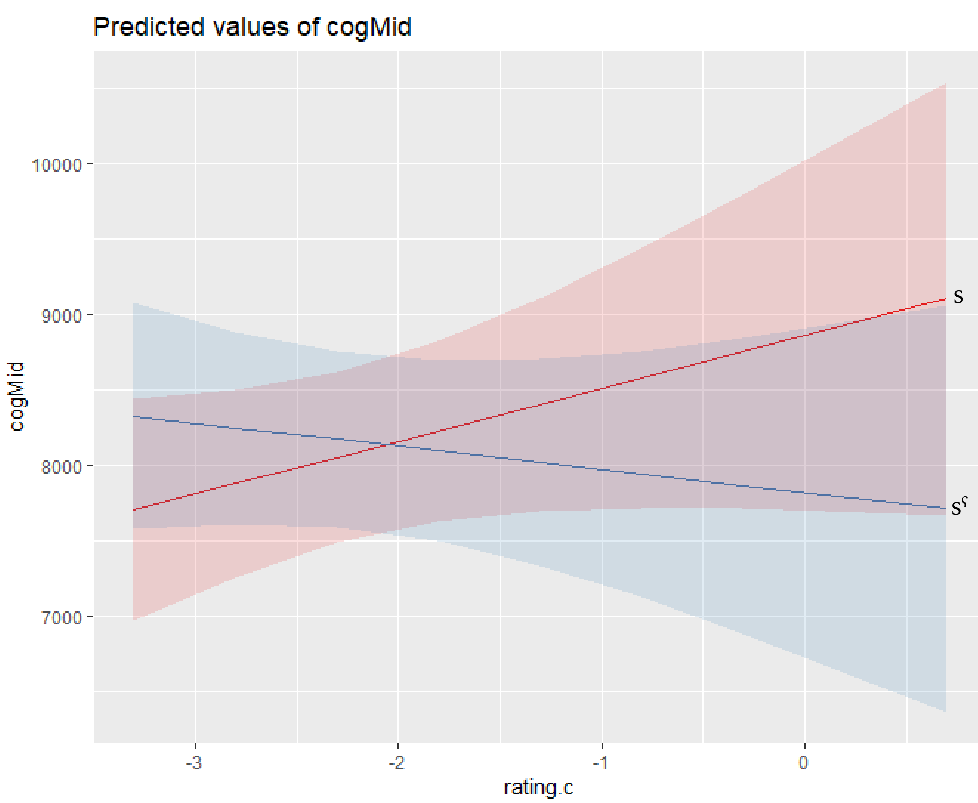


Figure 7.44 Interaction between fricative sound and foreign accent score for center of gravity

Standard deviation

FDHs' foreign accent score, LoR or religion did not have significant effects on fricative SD ($p>0.05$). Interaction between fricative sound and foreign accent score also proved to be insignificant.

Skewness

FDHs' foreign accent score, LoR or religion did not have significant effects on fricative skewness ($p>0.05$). Similar to CoG, there was a significant interaction effect between fricative sound and foreign accent score with regard to skewness. Figure 7.45 shows that the relationship between skewness and the type of fricative sound produced in FDH-directed speech depends considerably on FDHs' foreign accentedness score. Skewness of /s/ drops when the addressee is a FDH who is more native-like, while skewness of /sˤ/ increases when the addressee is a FDH who is more native-like. This indicates that NSs /s/ is more fronted when addressing FDHs' who are more native-like, while their /sˤ/ is more retracted when addressing FDHs' who are more native-like.

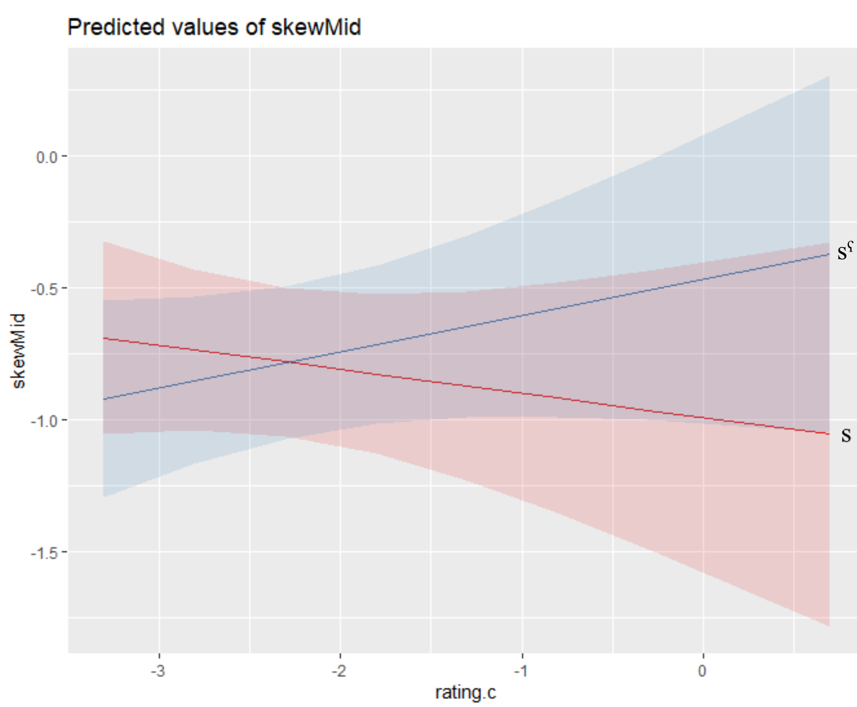


Figure 7.45 Interaction between fricative sound and foreign accent score for skewness

Kurtosis

FDHs' foreign accent score proved to be a significant predictor of fricative kurtosis. However, LoR, religion and fricative sound × foreign accent interaction did not prove significant. Figure 7.46 provides further evidence for the significant effect of foreign accent score on fricative kurtosis. The more native-like a FDH is, the higher fricative kurtosis is. Note that both /s/ and /sˤ/ take a similar trend in this effect. Higher kurtosis values with the more native -like group suggest that both fricative sounds are produced with a more retracted tongue.

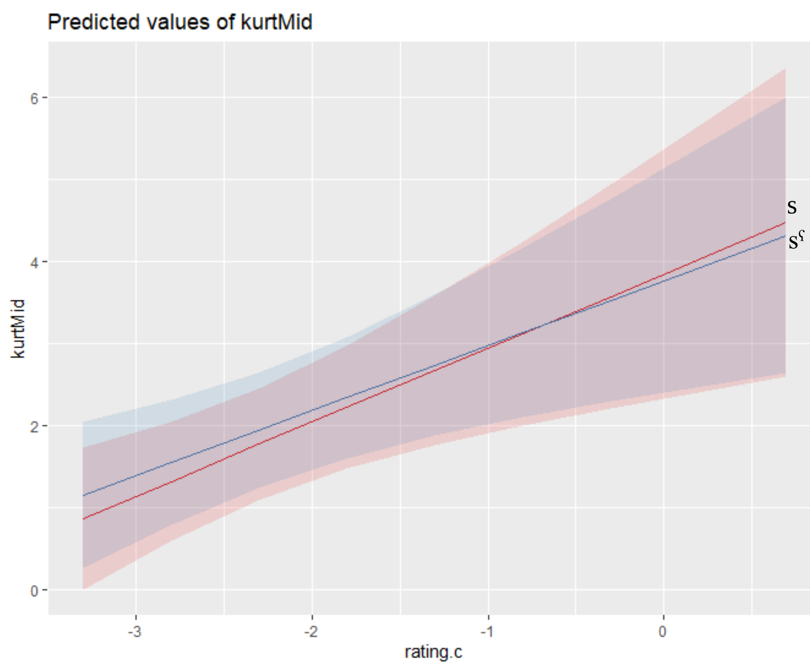


Figure 7.46 The effect of foreign accent rating on fricative kurtosis

7.4.2.2 Duration

Figure 7.47 and Table 7.17 show that duration of both /s/ and /sˤ/ is shorter in FDH-directed speech than in ADS (see Emmean); this proved to be significant for /sˤ/ only ($p<0.01$). The pattern in ADS and FDH-directed speech with regard to the difference in duration between /s/ and /sˤ/ was not entirely parallel. In ADS, /sˤ/ was 2.51 ms longer than /s/, whereas in FDS, /sˤ/ was 3.79 ms shorter than /s/. This was insignificant based on sound × speech style interaction revealed by the LMEM.

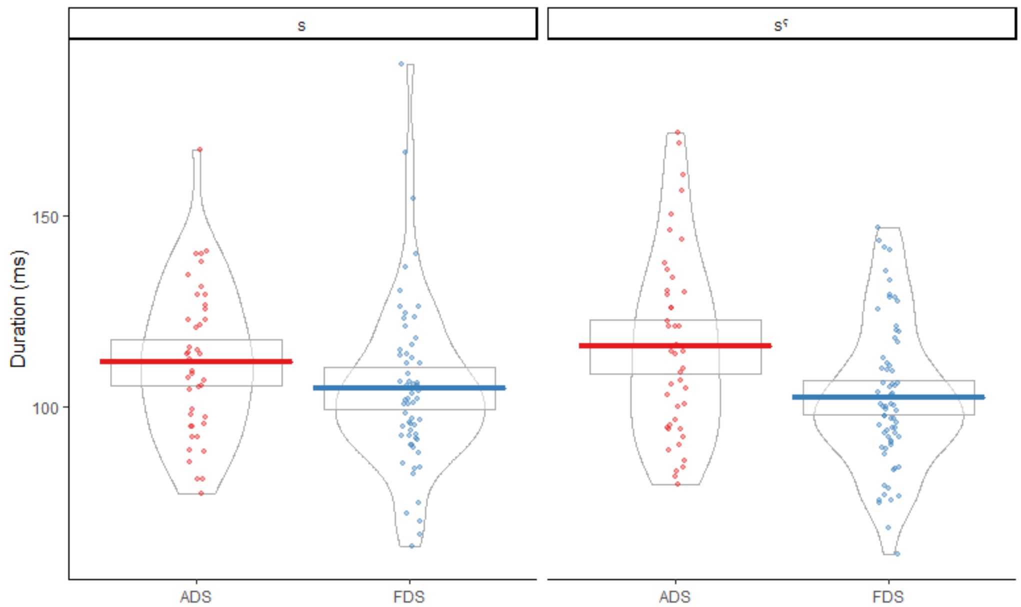


Figure 7.47 Duration of /s/ and /sˤ/ in FDH-directed speech and ADS

Sound	Register	emmean	SE	df	lower.CL	upper.CL
sˤ	ADS	114	3.79	59.0	106.8	122
	FDS	103	3.48	42.9	95.9	110
s	ADS	112	3.81	60.5	104.3	120
	FDS	107	3.54	45.4	99.6	114
Level	Contrasts	estimate	SE	df	t.ratio	p.value
sˤ	ADS-FDS	11.48	3.40	200	3.37	**0.0009**
s	ADS-FDS	5.19	3.55	202	1.462	0.14
ADS	sˤ -s	2.51	3.80	201	0.65	0.51
FDS	sˤ -s	-3.79	3.15	203	-1.203	0.23

Table 7.17 Results of *post hoc* pairwise comparison tests for the effect of speech style on duration of /s/ and /sˤ/

7.4.2.2.1 The effect of FDHs' foreign accented ratin, LoR, and religion on fricative duration

FDHs' foreign accent score, LoR or religion did not have significant effects on fricative duration ($p>0.05$; Table 7.18). This suggests that NSs did not vary the duration of their fricatives based on whether the FDH was native-/non-native-like, had a shorter/longer LoR or was Muslim/non-Muslim.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	107.69	10.89	36.16	9.88	
Target Sound:sˤ	2.45	7.06	113.9	0.34	0.72
Accent rating	-0.23	3.09	37.75	-0.07	0.94
LoR	0.02	0.58	24.6	0.04	0.96
Religion: Non-Muslim	-3.97	8.22	23.06	-0.48	0.63
Target Sound*Accent rating	3.05	2.64	114.84	1.15	0.25
Mixed effects: speaker	Var 141.2	SD 11.88			

Table 7.18 Results of LMEMs on the effect of foreign accentedness score, LoR and religion on fricative duration

7.4.2.3 Intensity

Figure 7.48 and Table 7.19 demonstrate that intensity of both /s/ and /sˤ/ is higher in FDH-directed speech than in ADS (see Emmean), though this was insignificant based on the LMEM and the *post hoc* tests. Both ADS and FDH-directed speech had a parallel pattern with regard to the pattern of intensity, with /sˤ/ having slightly lower intensity than /s/ in both speech styles.

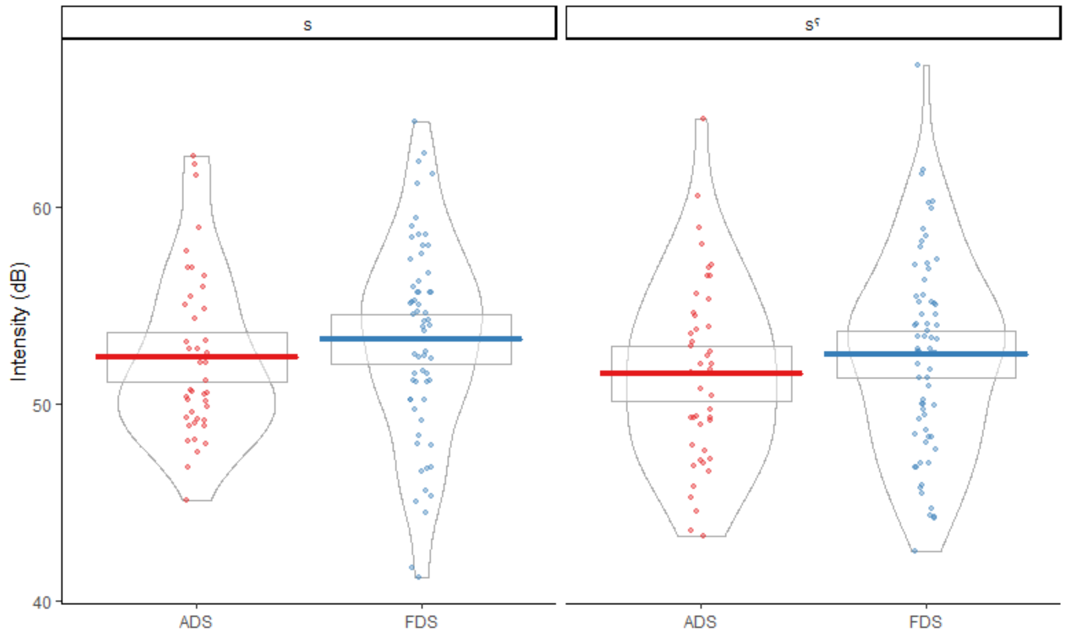


Figure 7.48 Intensity values of /s/ and /sˤ/ in FDH-directed speech and ADS

Sound	Register	emmean	SE	df	lower.CL	upper.CL
sˤ	ADS	51.9	0.902	53.8	50.1	53.7
	FDS	52.8	0.83	40.1	51.1	54.5
s	ADS	52.7	0.905	54.9	50.9	54.5
	FDS	54.5	0.84	42.3	51.3	54.7
Level	Contrasts	estimate	SE	df	t.ratio	p.value
sˤ	ADS-FDS	-0.87	0.77	200	-1.13	0.25
s	ADS-FDS	-0.31	0.804	201	-0.387	0.69
ADS	sˤ -s	-0.809	0.86	200	-0.93	0.34
FDS	sˤ -s	-0.24	0.71	202	-0.33	0.73

Table 7.19 Results of post hoc pairwise comparisons test for intensity of /s/ and /sˤ/

7.4.2.3.1 The effect of FDHs' foreign accented rating, LoR, and religion on fricative intensity

FDHs' foreign accent score, LoR or religion did not have significant effects on fricative intensity ($p>0.05$; Table 7.20). This suggests that NSs did not vary the intensity of their fricatives based on whether the FDH was native-/non-native -like, had a shorter/longer LoR or was Muslim/non-Muslim.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	56.5	2.52	34.01	22.33	
Target Sound:sˤ	0.12	1.84	112.12	0.06	0.94
Accent rating	0.75	0.72	36.15	1.04	0.3
LoR	0.11	0.13	20.23	0.87	0.39
Religion: Non-Muslim	-1.22	1.83	18.21	-0.66	0.51
Target Sound*Accent rating	0.22	0.69	113.56	0.33	0.74
Mixed effects: speaker	Var 6.07	SD 2.46			

Table 7.20 Results of LMEMs on the effect of foreign accentedness score, LoR and religion on fricative intensity

7.4.2.5 Other effects

Previously, we learned that there was a main effect of speech style on two metrics (skewness, and duration). To rule out any effects of condition order (whether FDH-directed speech/ADS was produced first or second in turn) on significant changes in these metrics, LMEM results in Table 7.21 show the difference between ADS and FDH-directed speech in condition order 1 (O1) and condition order 2 (O2).

Metric	Order	Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
/s/							
Skewness	O1	ADS (baseline)	-0.77	0.17	40.55	-4.406	
	FDS/ADS	FDS	0.33	0.21	50.2002	1.603	0.11
	O2	ADS (baseline)	-1.26	0.21	18.95	-5.86	
	ADS/FDS	FDS	0.51	0.209	43.04	2.47	**0.012 ***
Random Effects (O1)	Speaker	Variance 0.02	Std. Deviation 0.16				
Random Effects (O2)	Speaker	Variance 0.27	Std. Deviation 0.52				
/sˤ/							
Skewness	O1	ADS (baseline)	-0.76	0.1	29.34	-4.15	
	FDS/ADS	FDS	0.21	0.19	48.101	1.09	0.28
	O2	ADS (baseline)	-0.98	0.17	19.53	-5.68	
	ADS/FDS	FDS	0.404	0.17	49.52	2.27	**0.027 ***
Random Effects	Speaker	Variance 0.06	Std. Deviation 0.26				
Random Effects	Speaker	Variance 0.14	Std. Deviation 0.37				
Duration	O1	ADS (baseline)	109.13	5.03	25.43	21.66	
	FDS/ADS	FDS	-10.509	5.08	46.84	-2.06	**0.04 ***
	O2	ADS (baseline)	119.03	5.24	18.26	22.707	
	ADS/FDS	FDS	-12.36	4.803	49.39	-2.57	**0.01 ***
Random Effects (O1)		Variance 77.35	Std. Deviation 8.79				
Random Effects (O2)	Speaker	Variance 166.3	Std. Deviation 12.9				

Table 7.21 Results of LMEM models for the effect of condition order with regard to skewness and duration

The pattern of change in skewness of both /s/ and /sˤ/ was parallel in both condition orders. FDH-directed speech had constantly higher skewness regardless of whether the FDH carried out the task first or second in turn (Figures 7.49 and 7.50). This result rules out any effect of condition order on the change in skewness and confirms that speech style (ADS vs. FDH-directed speech) determines the noticeable adaptation in skewness of both continuants. It is important to highlight though that the difference in skewness of both /s/ and /sˤ/ between FDH-directed speech and ADS reached significance only in the second condition order where ADS was produced first in turn.

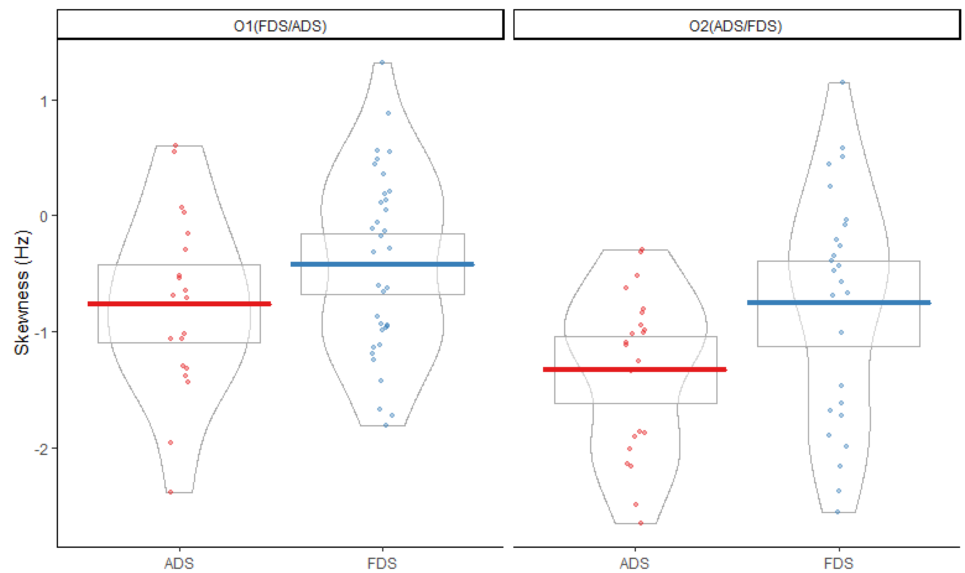


Figure 7.49 Skewness for /s/ in relation to condition order

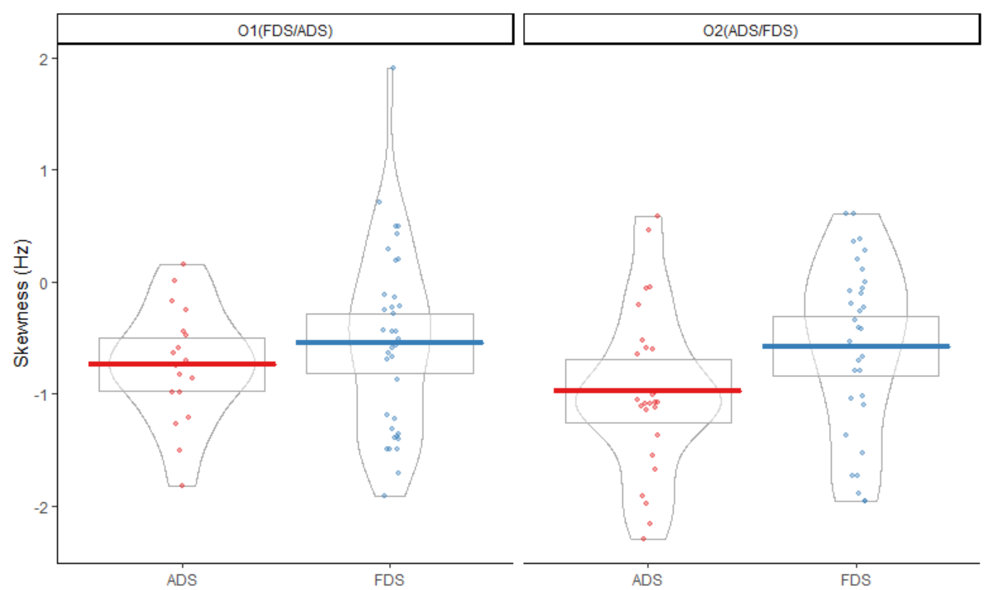


Figure 7.50 Skewness for /sˤ/ in relation to condition order

For the duration of /sˤ/, condition order did not have a significant effect on duration as /sˤ/ of FDH-directed speech was significantly shorter than that of ADS in both condition orders ($p<0.05$) (Table 7.22 and figure 7.51).

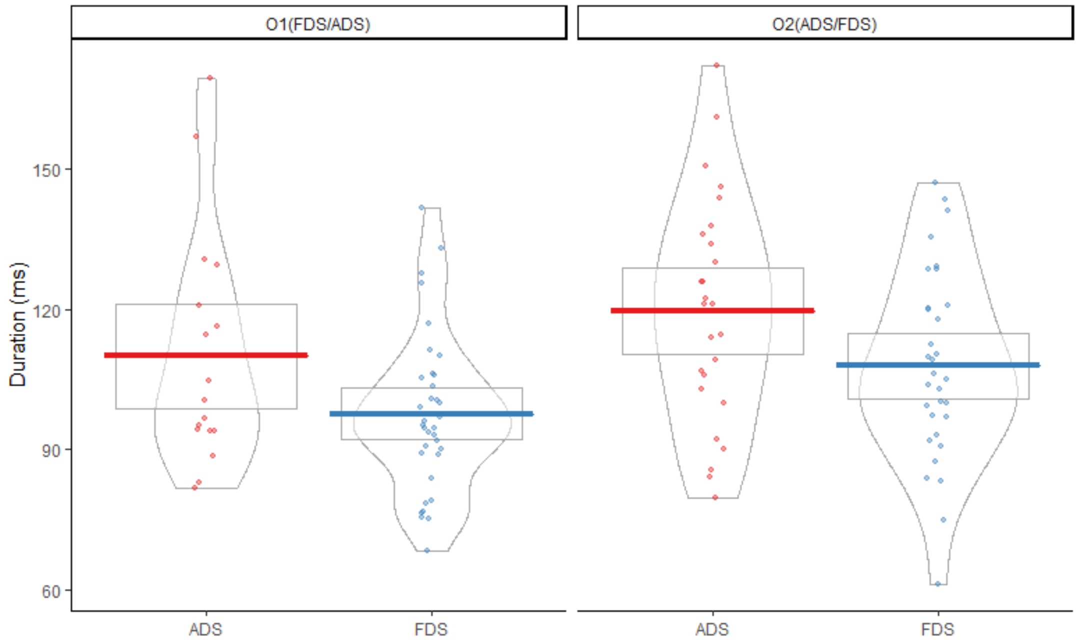


Figure 7.51 Duration for /sˤ/ in relation to condition order

Also, to rule out any effect of phrase boundary on the realization of the target consonants, likelihood ratio tests revealed that the full model containing speech style is always more significant than the reduced model. Hence, this suggests that the significant differences between FDH-directed speech and ADS for /s/ and /sˤ/ with regard to skewness and duration are not the by-product of whether the consonant appeared at a phrase boundary or not (Table 7.22)

Metric	**Model**	**Df**	**AIC**	**BIC**	**loglik**	**deviance**	**Chisq**	**Df**	**Pr(>Chisq)**
/s/									
Skewness	Reduced Model	4	260.74	271.32	-126.37	252.74			
	Full model	5	254.75	267.97	-122.37	244.75	7.99	1	**0.004 ****
/sˤ/									
Skewness	Reduced Model	4	248.18	259.05	-120.09	240.18			
	Full Model	5	243.38	256.98	-116.69	233.38	6.797	1	**0.009 ****
Duration	Reduced Model	4	999.95	1010.8	-495.98	991.95			
	Full Model	5	991.34	1004.9	-490.67	981.34	10.615	1	**0.001 ****

Table 7.22 Likelihood ratio tests for the effect of phrase boundary on skewness for /s/ and /sˤ/and duration for /sˤ/

A likelihood ratio test revealed that the difference between FDH-directed speech and ADS with regard to skewness of /s/ was not the by-product of word stress (whether /s/ appeared in a stressed syllable or not) as the full model containing speech style contributed significantly to the existing difference (Table 7.23).

Metric	Model	Df	AIC	BIC	loglik	deviance	Chisq	Df	Pr(>Chisq)
/s/									
Skewness	Reduced Model	4	262.18	272.76	-127.09	254.18			
	Full Model	5	256.18	269.40	-123.09	246.18	7.99	1	**0.004 ****

Table 7.23 Likelihood ratio test for the effect of word stress on skewness for /s/

7.4.2.6 Absolute formants of following vowels

LMEMs showed that F1 and F2 at the onset of the following vowels were significant predictors of emphasis as shown in Table 7.24 ($p<0.01$). However, there was no significant effect of sound type on F3 of the following vowels ($p>0.05$). Figure 7.52 and Table 7.25 show that the differences between ADS and FDH-directed speech with regard to F1, F2 and F3 of the vowels following both /s/ and /sˤ/ are negligible (see Emmean). The direction of F1 and F2 from emphatic to plain was the same in both FDH-directed speech and ADS. F1 of vowels in an emphatic environment was significantly higher than that in a plain environment in both ADS and FDH-directed speech. F2 of vowels in an emphatic environment was significantly lower than that in a plain environment in both ADS and FDH-directed speech. The direction of F3 from emphatic to plain was slightly different in ADS compared to FDH-directed speech as Table 7.25 shows, but this was insignificant.

Metric	**Predictor**	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
F1 (Bark)	(Intercept)	6.02	0.06	81.81	86.63	
	Sound /s/	-0.9003	0.07	201.305	-11.29	**<0.0001*****
	Speech FDS	0.01	0.07	199.99	0.19	0.84
	/s/: FDS	0.08	0.103	201.34	0.83	0.403
Random effects	speaker	Variance 0.03	Std.Deviation 0.19			
F2 (Bark)	(Intercept)	9.93	0.08	86.46	115.802	
	Sound /s/	2.906	0.103	200.34	28.15	**<0.0001*****
	Speech FDS	0.15	0.09	198.65	1.64	0.103
	/s/: FDS	-0.107	0.13	200.52	-0.808	0.42
Random effects	speaker	Variance 0.04	Std.Deviation 0.21			
F3 (Bark)	(Intercept)	15.74	0.09	54.41	167.71	
	Sound /s/	0.07	0.09	197.53	0.85	0.39
	Speech FDS	0.03	0.08	196.52	0.43	0.66
	/s/: FDS	-0.08	0.12	197.44	-0.73	0.46
Random effects	speaker	Variance 0.18	Std.Deviation 0.42			

Table 7.24 Results of LMEMs for the effect of speech style on formant frequencies of vowels following plain/emphatic consonants

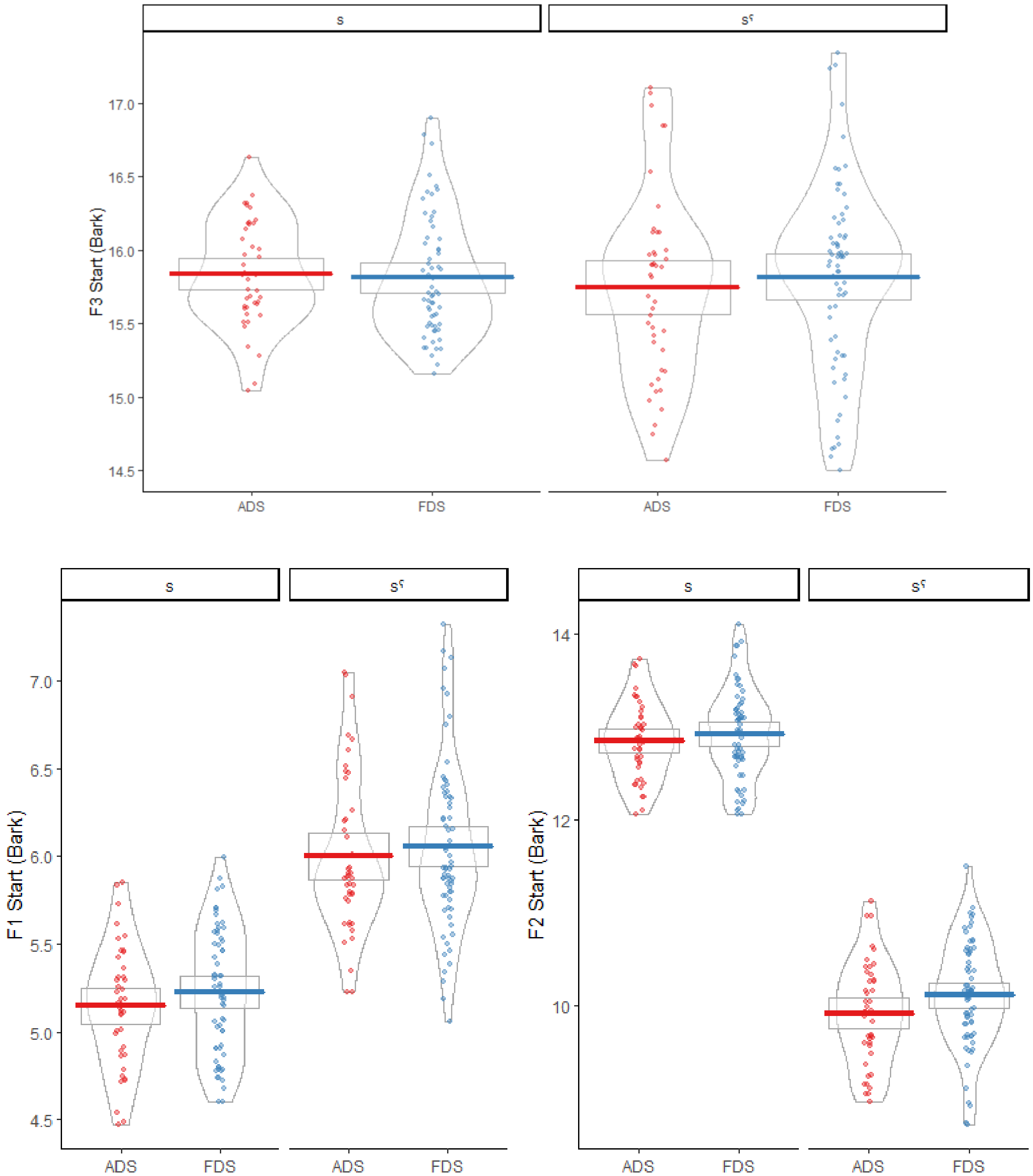


Figure 7.52 Absolute formant frequencies of vowels following /s/ and /sˤ/ in ADS and FDH-directed speech

Metric	Sound	Register	emmean	SE	df	lower.CL	upper.CL
F1	$s^ʕ$	ADS	6.03	0.07	81.3	5.89	6.17
		FDS	6.04	0.06	54.3	5.92	6.17
	s	ADS	5.13	0.07	85.0	4.99	5.27
		FDS	5.23	0.06	59.0	5.10	5.36
	Level	**Contrasts**	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	$s^ʕ$	ADS-FDS	-0.01	0.07	202	-0.19	0.84
	s	ADS-FDS	-0.1003	0.07	205	-1.33	0.18
	ADS	$s^ʕ$ -s	0.900	0.08	203	11.18	**<.0001**
	FDS	$s^ʕ$ -s	0.81	0.06	206	12.28	**<.0001**
	Sound	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
F2	$s^ʕ$	ADS	9.94	0.08	95.7	9.77	10.1
		FDS	10.09	0.07	62.3	9.94	10.2
	s	ADS	12.85	0.08	101.0	12.67	13.0
		FDS	12.89	0.07	67.9	12.73	13.0
	Level	**Contrasts**	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	$s^ʕ$	ADS-FDS	-0.15	0.09	203	-1.62	0.105
	s	ADS-FDS	-0.04	0.09	206	-0.44	0.65
	ADS	$s^ʕ$ -s	-2.91	0.104	205	-27.87	**<.0001**
	FDS	$s^ʕ$ -s	-2.80	0.08	207	-32.63	**<.0001**
	Sound	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
F3	$s^ʕ$	ADS	15.7	0.09	58.1	15.6	15.9
		FDS	15.8	0.08	42.1	15.6	16.0
	s	ADS	15.8	0.09	59.5	15.6	16.0
		FDS	15.8	0.08	44.9	15.6	16.0
	Level	**Contrasts**	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	$s^ʕ$	ADS-FDS	-0.03	0.08	200	-0.43	0.66
	s	ADS-FDS	0.05	0.08	202	0.58	0.55
	ADS	$s^ʕ$ -s	-0.07	0.09	201	-0.84	0.4007
	FDS	$s^ʕ$ -s	0.008	0.07	202	0.11	0.909

Table 7.25 Results of post hoc pairwise comparisons test for absolute formants of vowels following /s/ and /$s^ʕ$/

7.4.2.6.1 The effect of FDHs' foreign accented rating, LoR, and religion on vowel formant frequencies

FDHs' foreign accent score, LoR and religion did not have any significant effects on vowel F1 following fricatives as shown in Table 7.26 ($p>0.05$). Religion, however, was a significant predictor for changes in vowel F2 ($p<0.05$). F2 values of vowels following fricatives were significantly lower when the addressees were non-Muslim FDHs. This trend is similar in both fricative sounds as shown in Figure 7.53. This suggests that both fricatives were more retracted in NSs' productions when addressing non-Muslim FDHs. To this end, this finding does not indicate expansion of the contrast in NSs' production.

Metric	Predictor	Estimate	Std.Error	df	t-value	
	Intercept	5.16	0.07	26.79	67.64	
	Target Sound:sˤ	0.81	0.06	115.97	12.29	<2e-16 ***
F1	Accent rating	-0.04	0.06	38.05	-0.76	0.45
	LoR	-0.01	0.01	21.79	-0.97	0.34
	Religion: Non-Muslim	0.19	0.15	19.83	1.25	0.22
	Target Sound*Accent rating	0.07	0.05	113.12	1.25	0.21
Mixed effects	speaker	Var 0.04	SD 0.209			
	Intercept	12.98	0.07	30.13	163.58	
	Target sound:sˤ	-2.78	0.08	123.26	31.25	<2e-16 ***
F2	Accent rating	0.11	0.06	50.54	1.65	0.1
	LoR	0.01	0.01	21.07	1.69	0.1
	Religion: Non-Muslim	-0.33	0.15	16.58	-2.25	0.0381 *
	Target Sound*Accent rating	-0.1	0.07	118.24	-1.34	0.18
Mixed effects	speaker	Var 0.02	SD 0.14			

Table 7.26 LMEMs results for the effect of foreign accent, LoR and religion on vowel F1 and F2

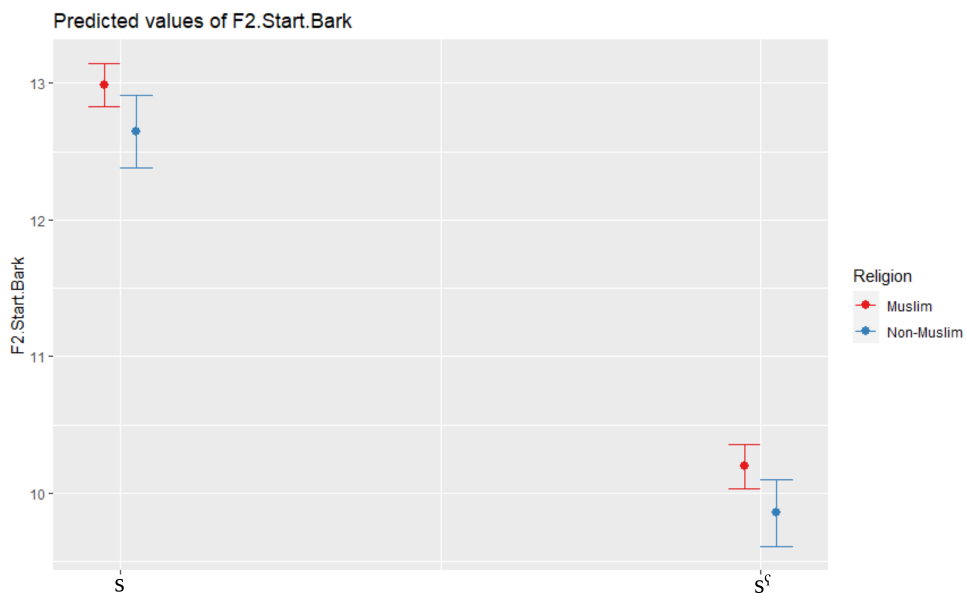


Figure 7.53 The effect of religion on vowel F2 following fricative sounds

7.4.3 Plain/emphatic stop consonants

Unlike fricatives, LMEMs revealed a main effect of sound type on M1, M2 and M3 but not M4 of the stop consonants as Table 7.27 shows (p<0.01). Moreover, a significant effect of sound type was evident with regard to VOT and intensity (p<0.01). Longer VOT in /t/ suggests that this sound is aspirated unlike its emphatic counterpart (Al-Ani 1970; Giannini and Pettorino 1982; Al-Nuzaili 1993). This suggests that these cues are likely enhancement cues for the plain/emphatic contrast in FDH-directed speech. However, there was a main effect of speech style on one metric only, that was intensity (p<0.01). Interaction between sound type and speech style yielded non-significant values in all metrics (p>0.05). Subsequent *post hoc* results for each metric will be provided in the following sections.

Metric	Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
Centre of Gravity	(Intercept)	1107.0	248.1	100.5	4.46	
	Sound: /t/	1234.6	309.8	155.9	3.98	**0.0001*** **
	Speech: FDS	-229.6	272.1	160.7	-0.84	0.4
	/t/*FDS	195.3	395.1	158.3	0.49	0.62
Random effects	speaker	Variance 310729	Std.Deviation 557.4			
Sd. Deviation	(Intercept)	1805.9	236.7	69.1	7.63	
	Sound: /t/	893.7	256.9	154.3	3.47	**0.0006 *** **
	Speech: FDS	-186.7	226.6	157.0	-0.82	0.41
	/t/*FDS	128.1	328.4	155.9	0.39	0.69
Random effects	speaker	Variance 508326	Std.Deviation 713			
Skewness	(Intercept)	5.34	0.64	86.01	8.26	
	Sound: /t/	-2.79	0.76	155.15	-3.67	**0.0003*** **
	Speech: FDS	0.81	0.67	58.86	1.21	0.22
	/t/*FDS	-0.31	0.97	157.25	-0.32	0.74
Random effects	speaker	Variance 2.84	Std.Deviation 1.68			
Kurtosis	(Intercept)	36.07	10.67	104.96	3.38	
	Sound /t/	-23.59	13.34	157.36	-1.76	0.07
	Speech: FDS	39.55	11.72	161.8	3.37	**0.0009 *** **
	/t/*FDS	-30.76	17.01	159.66	-1.809	0.078
Random effects	speaker	Variance 569.8	Std.Deviation 23.87			
VOT	(Intercept)	18.82	1.58	93.05	11.87	
	Sound /t/	5.806	1.87	157.38	3.09	**0.002** **
	Speech FDS	-2.58	1.65	160.72	-1.56	0.11
	/t/*FDS	1.72	2.39	159.26	0.71	0.47
Random effects	speaker	Variance 16.87	Std.Deviation 4.107			
Intensity	(Intercept)	48.07	0.93	67.75	51.32	
	Sound /t/	-2.23	1.007	154.28	-2.21	**0.02 ***
	Speech FDS	2.89	0.88	156.91	3.26	**0.001** **
	/t/*FDS	-1.75	1.28	155.91	-1.36	0.17
Random effects	speaker	Variance 8.17	Std.Deviation 2.85			

Table 7.27 Results of LMEMs for the effect of speech style on acoustic properties of plain/emphatic stops

7.4.3.1 Spectral moments

- Center of Gravity: CoG of both /t/ and /tˤ/ was, generally, slightly higher in ADS than FDH-directed speech, though this was not significant (see Table 7.28 and Figure 7.54).
- Standard Deviation: Like the pattern of CoG, SD of both /t/ and /tˤ/ was higher in ADS than in FDH-directed speech as shown in Table 7.28 and Figure 7.54. The difference between FDH-directed speech and ADS with regard to SD was insignificant though.
- Skewness: Skewness of both /t/ and /tˤ/ was lower in ADS than in FDH-directed speech as Table 7.28 and Figure 7.54 show. The difference between FDH-directed speech and ADS with regard to skewness was insignificant though.
- Kurtosis: kurtosis of both /t/ and /tˤ/ was lower in ADS compared to FDH-directed speech, but this was insignificant as demonstrated in Table 7.28 and Figure 7.54. An interesting result revealed by *post hoc* analyses was that there was a significant difference between /t/ and /tˤ/ with regard to kurtosis in FDH-directed speech but not in ADS ($p<0.01$). Although there was no statistical difference between /t/ and /tˤ/ with regard to kurtosis in ADS, the trend of kurtosis in both ADS and FDS showed an increase from plain to emphatic.

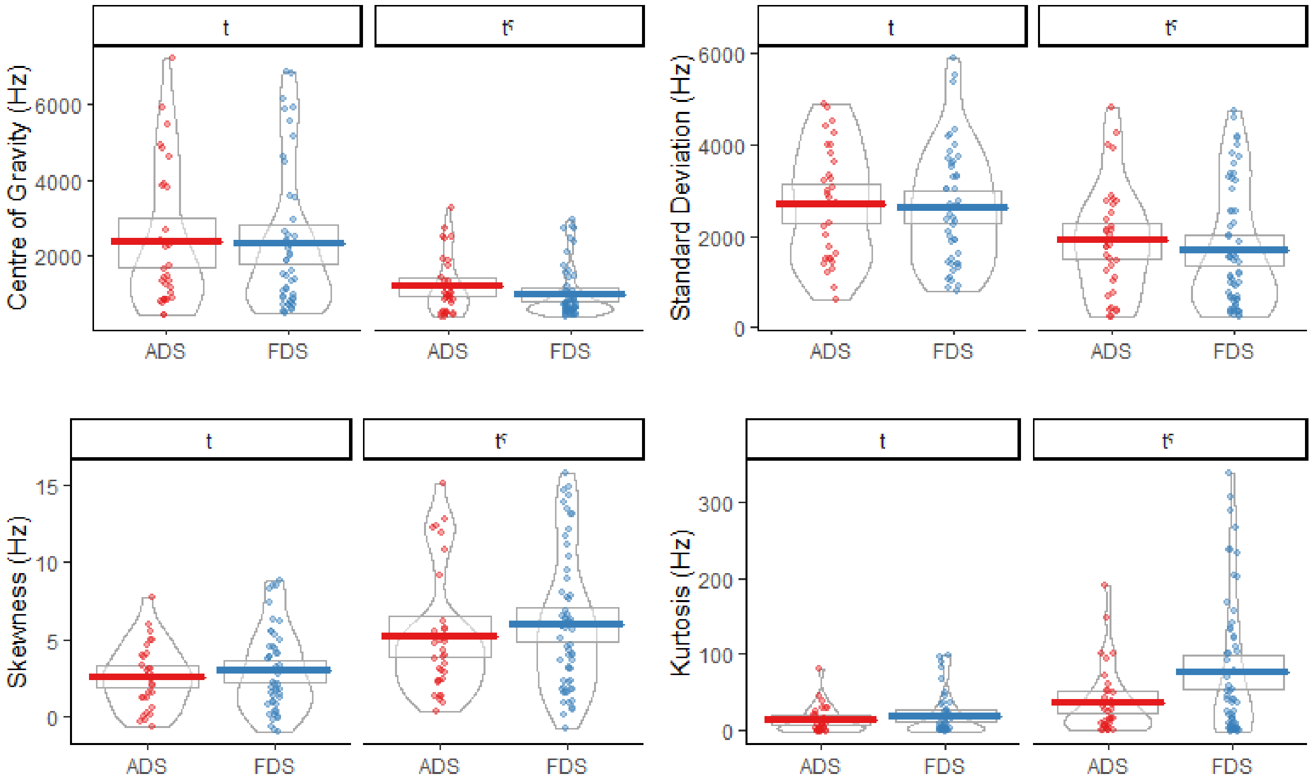


Figure 7.54 Spectral moments of /t/ and /tˤ/ in FDH-directed speech and ADS

Metric	**Sound**	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Center of Gravity	tˤ	ADS	1107	253	104.8	605	1608
		FDS	877	211	60.4	456	1299
	t	ADS	2342	259	114.8	1829	2854
		FDS	2307	226	73.5	1856	2758
	Level	Contrasts	estimate	SE	df	t.ratio	p.value
	tˤ	ADS-FDS	229.6	276	164	0.83	0.406
	t	ADS-FDS	34.3	297	165	0.116	0.908
	ADS	tˤ -t	-1235	313	159	-3.94	**0.0001**
	FDS	tˤ -t	-1430	254	169	-5.636	**<.0001**
Standard Deviation	**Sound**	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
	tˤ	ADS	1806	241	70.5	1325	2287
		FDS	1619	212	44.4	1193	2046
	t	ADS	2700	245	75.4	2211	3188
		FDS	2641	223	52.6	2193	3089
	Level	**Contrasts**	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	tˤ	ADS-FDS	186.7	229	160	0.814	0.41
	t	ADS-FDS	58.6	247	161	0.23	0.81
	ADS	tˤ -t	-894	260	157	-3.441	**0.0007**
	FDS	tˤ -t	-1022	212	163	-4.82	**<.0001**
	Sound	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Skewness	tˤ	ADS	5.34	0.65	88.5	4.03	6.65
		FDS	6.15	0.56	52.5	5.03	7.28
	t	ADS	2.54	0.67	96.0	1.21	3.88
		FDS	3.04	0.59	63.3	1.84	4.24
	Level	**Contrasts**	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	tˤ	ADS-FDS	-0.81	0.68	162	-1.19	0.23
	t	ADS-FDS	-0.49	0.73	163	-0.67	0.49
	ADS	tˤ -t	2.80	0.77	158	3.62	**0.0004**
	FDS	tˤ -t	3.11	0.62	166	4.96	**<.0001**
	Sound	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
Kurtosis	tˤ	ADS	36.1	10.87	104.8	14.51	57.6
		FDS	75.6	9.05	60.3	57.52	93.7
	t	ADS	12.5	11.13	114.9	-9.56	34.5
		FDS	21.3	9.72	73.4	1.89	40.6
	Level	**Contrasts**	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	tˤ	ADS-FDS	-39.55	11.9	164	-3.33	**0.001**
	t	ADS-FDS	-8.79	12.8	165	-0.68	0.49
	ADS	tˤ -t	23.6	13.5	159	1.74	0.082
	FDS	tˤ -t	54.4	10.9	169	4.97	**<.0001**

Table 7.28 Results of post hoc pairwise tests for spectral moments of plain/emphatic stops

7.4.3.1.1 The effect of FDHs' foreign accented rating, LoR, and religion on spectral moments

LMEMs revealed that FDHs' foreign accent score, LoR and religion did not play any significant role in the spectral moments of the stop consonants produced by NSs. However, there was a significant sound x accent rating interaction with regard to skewness and kurtosis. These effects will be discussed in the following lines.

Metric	Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
Cog	Intercept	2217.32	295.3	28.2	7.5	
	Target Sound:tˤ	-1433.534	242.6	100.25	-5.9	**4.76e-08 *****
	Accent rating	44.55	178.4	28.08	0.25	0.8
	LoR	-8.5	38.34	24.97	-0.22	0.82
	Religion: Non-Muslim	818.94	568.81	27.71	1.44	0.16
	Target Sound*Accent rating	155.75	196.52	97.51	0.79	0.43
Mixed effects	speaker	Var 376969	SD 614			
SD	Intercept	2528.44	256.57	29.45	9.85	
	Target Sound:tˤ	-980.72	195.21	95.65	-5.02	**2.35e-06 *****
	Accent rating	59.86	189.72	26.97	0.31	0.75
	LoR	7.83	41.08	27.02	0.19	0.85
	Religion: Non-Muslim	499.48	599.93	29.58	0.83	0.41
	Target Sound*Accent rating	196.53	157.21	93.54	1.25	0.21
Mixed effects	speaker	Var 664410	SD 815.1			
Skew	Intercept	3.2	0.69	29.9	4.58	
	Target Sound: tˤ	2.85	0.61	97.42	4.66	**9.75e-06 *****
	Accent rating	-0.27	0.51	26.81	-0.54	0.59
	LoR	-0.05	0.11	25.7	-0.53	0.59
	Religion: Non-Muslim	-0.79	1.63	28.43	-0.48	0.63
	Target Sound*Accent rating	-0.99	0.49	94.81	-2.02	**0.04 ***
Mixed effects	speaker	Var 4.09	SD 2.02			
Kurt	Intercept	22.91	11.38	35.61	2.01	
	Target Sound: tˤ	50.69	12.14	102.88	4.17	**6.28e-05 *****
	Accent rating	-4.66	8.23	31.59	-0.56	0.57
	LoR	-0.18	1.75	26.3	-0.1	0.91

	Religion: Non-Muslim	-19.07	26.19	29.1	-0.72	0.47
	Target Sound*Accent rating	-27.93	9.86	100.62	-2.83	**0.005 ****
Mixed effects	speaker	Var 609.6	SD 24.69			

Table 7.29 LMEM results for the effect of foreign accent score, LoR and religion on stop spectral moments

Skewness

Figure 7.55 shows the interaction effect between the stop consonant and foreign accent score. There is a general decrease in skewness values for both stop consonants as FDHs foreign accent rating scores get closer to native-likeness. However, the skewness values drop sharply for /tˤ/ compared to those of /t/. Figure 7.56 shows that this trend is similar in both Muslim and non-Muslim FDHs. To this end, the significant interaction effect still does not indicate expansion of the contrast.

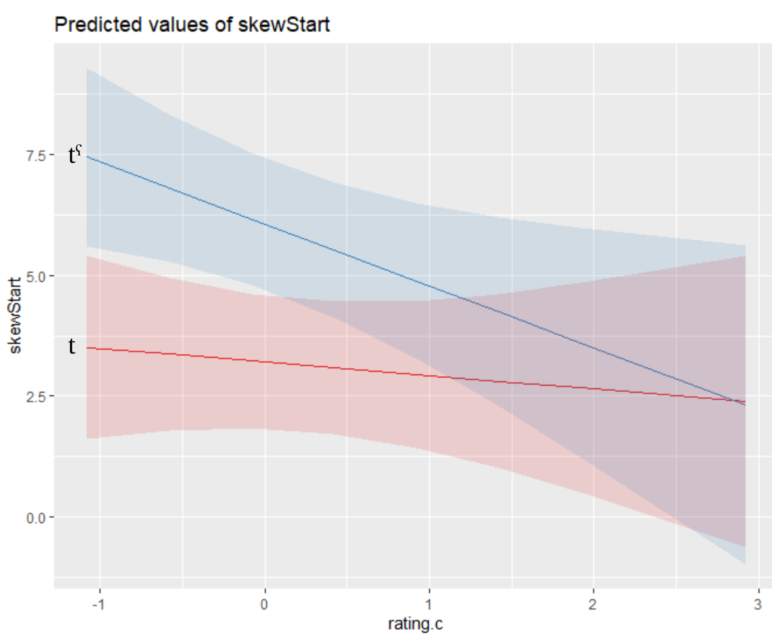


Figure 7.55 Interaction between stop sound and foreign accent score for skewness

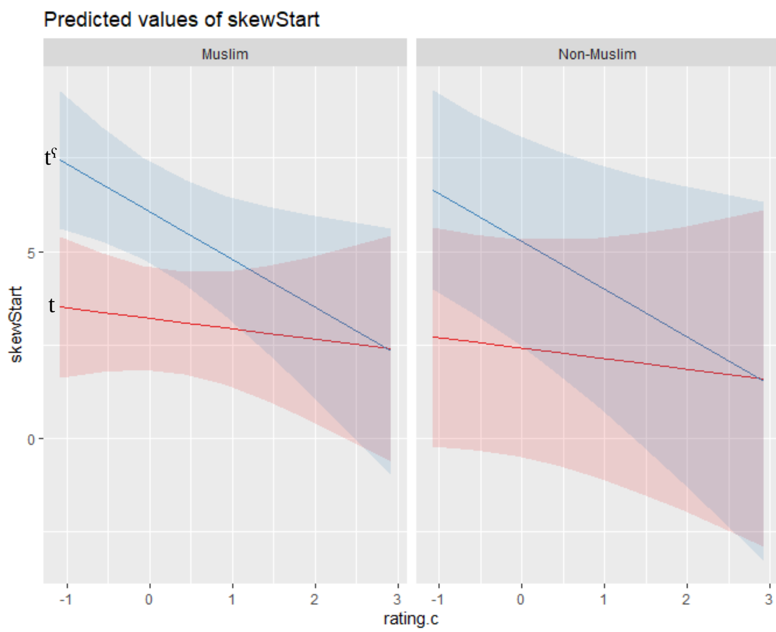


Figure 7.56 Interaction between stop sound and foreign accent score for skewness divided by religion group

Kurtosis

Figure 7.57 shows the interaction between stop sound and foreign accent rating scores. The trend is very similar to that of skewness. While kurtosis of /t/ drops slightly as FDHs' native-likeness increases, that of /tʕ/ drops sharply. As the two sounds still follow a similar trend, we cannot conclude that there is expansion in this contrast with regard to kurtosis. Figure 7.58 confirms that this interaction effect is similar for both Muslims and non-Muslims.

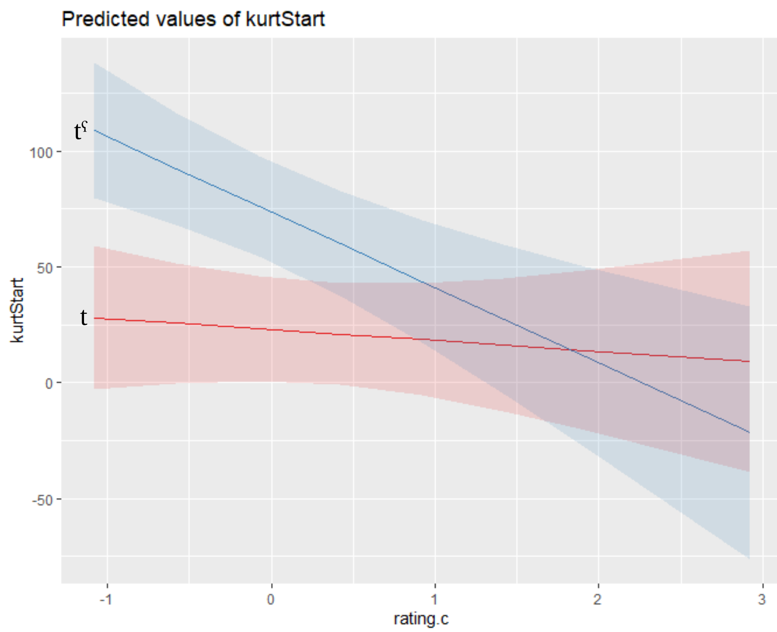


Figure 7.57 Interaction between stop sound and foreign accent score for kurtosis

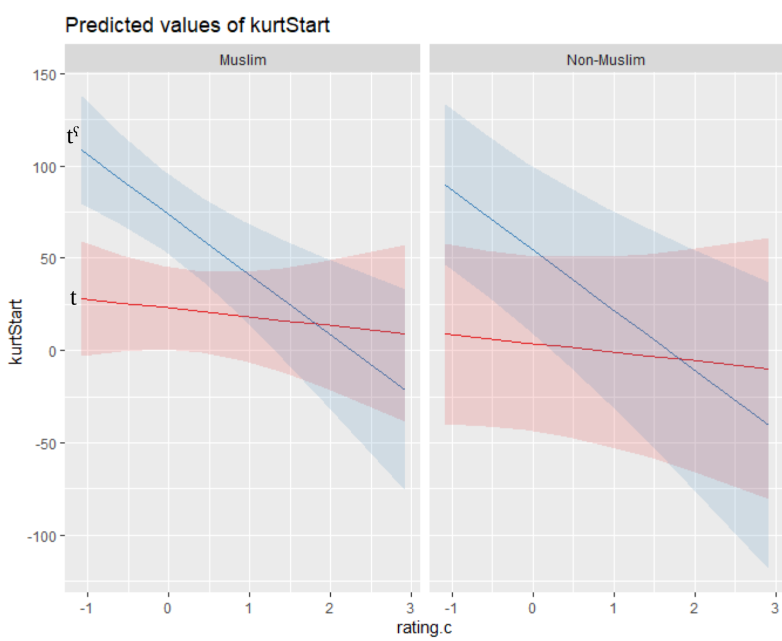


Figure 7.58 Interaction between stop sound and foreign accent score for kurtosis across Muslims and Non-Muslims

6.4.3.2 VOT

Generally, /t/ and /tˤ/ in FDH-directed speech had shorter VOTs than those in ADS, but this was very slight and statistically negligible (see Emmean in Table 7.30 and Figure 7.58). LMEM revealed a significant difference between /t/ and /tˤ/ in terms of VOT in both FDH-directed speech and ADS demonstrated in Table 7.30 ($p<0.01$).

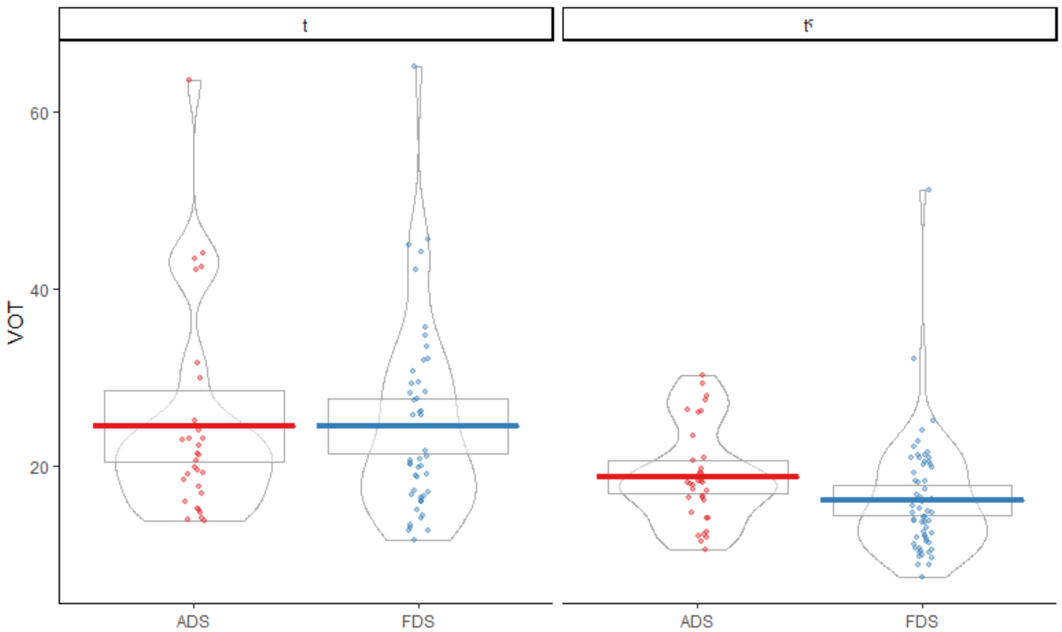


Figure 7.59 VOT values of /t/ and /tˤ/ in FDH-directed speech and ADS

Sound	Register	emmean	SE	df	lower.CL	upper.CL
tˤ	ADS	18.8	1.62	88.8	15.6	22.0
	FDS	16.2	1.38	52.5	13.5	19.0
t	ADS	24.6	1.65	96.4	21.4	27.9
	FDS	23.8	1.47	63.4	20.8	26.7
Level	Contrasts	estimate	SE	df	t.ratio	p.value
tˤ	ADS-FDS	2.58	1.67	162	1.54	0.12
t	ADS-FDS	0.86	1.80	163	0.48	0.63
ADS	tˤ -t	-5.81	1.90	158	-3.05	**0.002**
FDS	tˤ -t	-7.53	1.54	167	-4.87	**<.0001**

Table 7.30 Results of post hoc pairwise comparison tests for VOT of /t/ and /tˤ/

7.4.3.2.1 The effect of FDHs' foreign accented rating, LoR, and religion on stop duration

LMEM revealed that foreign accent ratings, LoR and religion did not play any significant effects on duration of stop consonants in FDH-directed speech (Table 7.31). This indicates that FDH-directed speech is not triggered by whether the FDHs' are foreign accented, have shorter LoR or are non-Muslims.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	24.74	1.91	32.05	12.89	
Target Sound:tˤ	-7.25	1.48	100.21	-4.89	3.73e-06 ***
Accent rating	1.006	1.16	31.17	0.86	0.39
LoR	-0.2	0.25	28.96	-0.83	0.41
Religion: Non-Muslim	1.88	3.69	31.96	0.51	0.61
Target Sound*Accent rating	0.04	1.19	97.84	0.03	0.97
Mixed effects: speaker	Var 18.45	SD 4.29			

Table 7.31 Results of LMEMs for the effect of foreign accent rating, LoR and religion on VOT

7.4.3.3 Intensity

Intensity was found to be significantly higher in an emphatic environment compared to a plain one in both FDH-directed speech and ADS based on LMEMs. LMEMs also demonstrated that there was a main effect of speech style on intensity ($p<0.01$). *Post hoc* tests revealed that the difference in intensity between FDH-directed speech and ADS was significant for /tˤ/ only as Table 7.32 and Figure 7.59 demonstrate. Intensity of /tˤ/ was 2.9 dB higher in FDH-directed speech than in ADS, whereas intensity of /t/ was 1.15 dB higher in FDH-directed speech than in ADS.

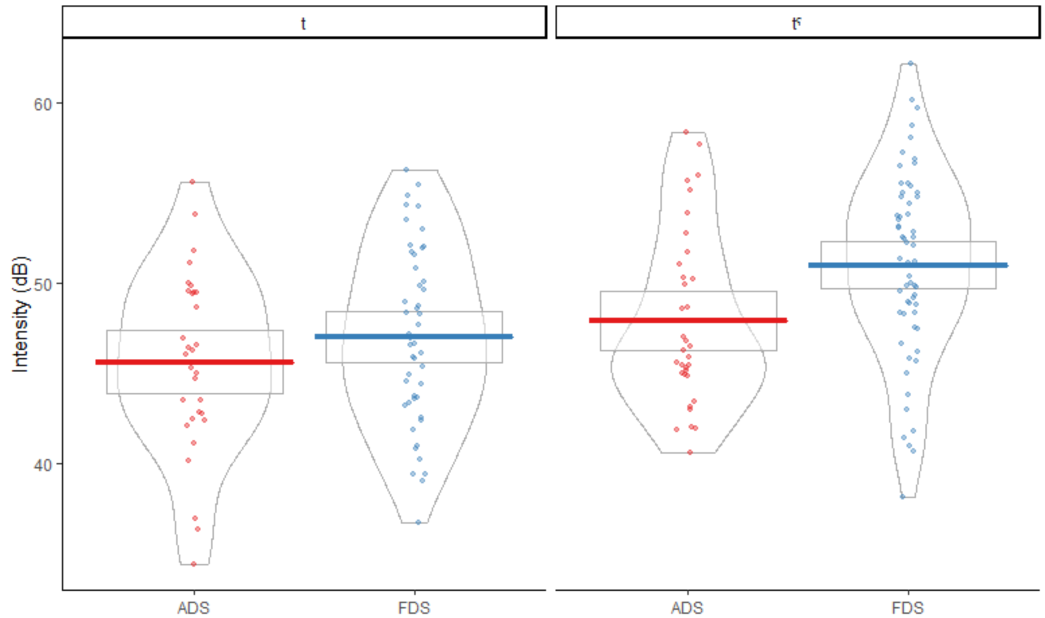


Figure 7.60 Intensity values of /t/ and /tˤ/ within and across speech style

Sound	Register	emmean	SE	df	lower.CL	upper.CL
tˤ	ADS	48.1	0.95	68.8	46.2	50.0
	FDS	51.0	0.84	43.7	49.3	52.7
t	ADS	45.8	0.97	73.5	43.9	47.8
	FDS	47.0	0.88	51.6	45.2	48.8
	Contrasts	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
tˤ	ADS-FDS	-2.90	0.89	159	-3.22	**0.001**
t	ADS-FDS	-1.15	0.96	160	-1.18	0.23
ADS	tˤ -t	2.24	1.01	157	2.19	**0.029**
FDS	tˤ -t	3.99	0.83	163	4.79	**<.0001**

Table 7.32 Results of post hoc pairwise comparison tests for intensity of /t/ and /tˤ/

7.4.3.3.1 The effect of FDHs' foreign accented rating, LoR, and religion on stop intensity

LMEM revealed that foreign accent rating did not have any significant effects on stop intensity in FDH-directed speech (Table 7.33). LoR has a significant effect on stop intensity as shown in table 7.33. Figure 7.60 shows that this effect is similar for both /t/ and /tˤ/. As FDHs LoR extends, stop intensity increases. Also, LMEM showed a significant effect of religion on stop intensity. As shown in Figure 7.60, stop intensity of /t/ is 4.94 dB higher than that of /t/ in speech addressed to non-Muslim FDHs. The target sound x accent rating interaction effect did not reach significance, though it had a nearly ceiling effect (p=0.05). Figure 7.61 demonstrates this interaction effect. It shows that intensity of /tˤ/ decreases sharply as FDHs' sound more native-like, while that of /t/ remains constant

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	47.67	0.94	32.01	50.6	
Target Sound:tˤ	3.76	0.81	98.18	4.61	1.19e-05 ***
Accent rating	-0.13	0.69	28.82	-0.19	0.84
LoR	0.34	0.14	27.76	2.32	0.02 *
Religion: Non-Muslim	-4.94	2.19	30.62	-2.25	0.03 *
Target Sound*Accent rating	-1.25	0.65	95.79	-1.91	0.05
Mixed effects: speaker	VAR 7.53	SD 2.74			

Table 7.33 LMEM results on the effect of FDHs' foreign accented rating, LoR, and religion on stop intensity

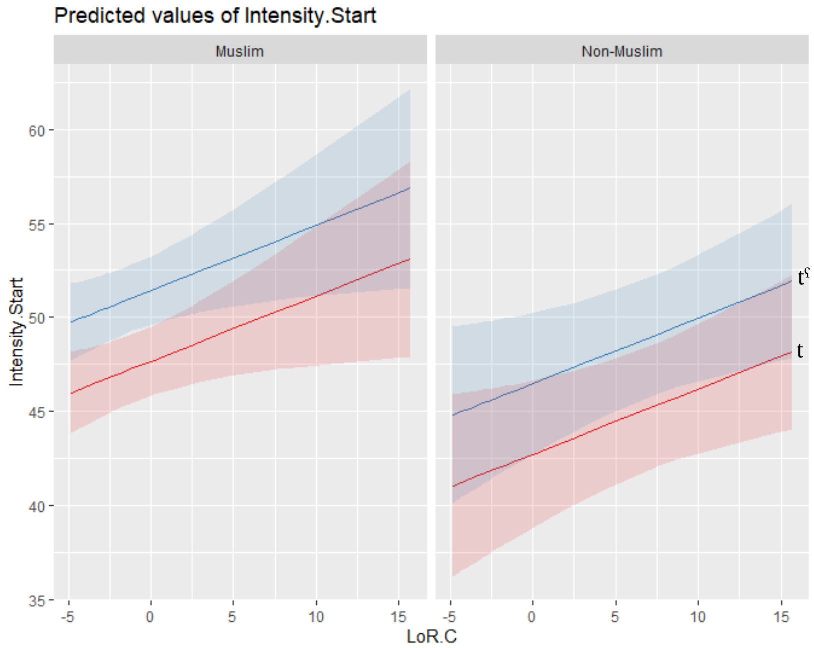


Figure 7.61 The effect of LoR on stop intensity in both Muslim and non-Muslim FDHs

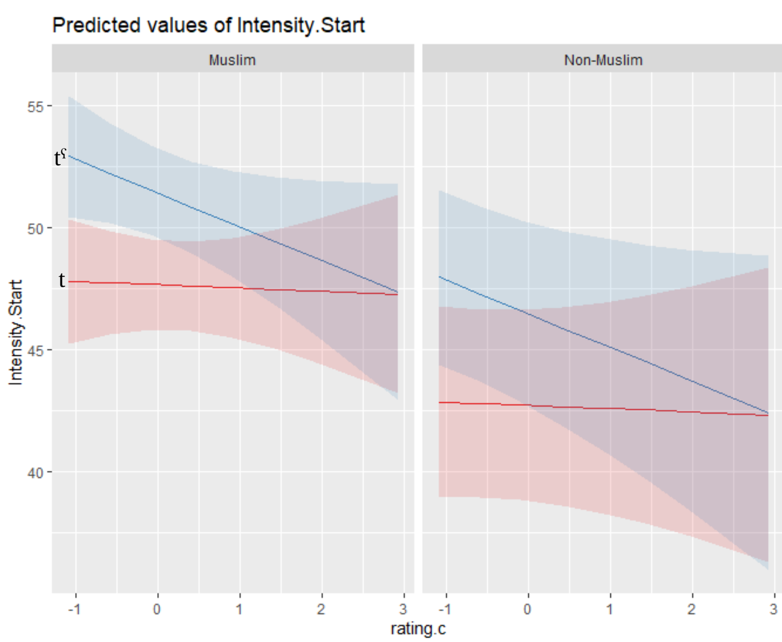


Figure 7.62 The effect of target sound and foreign accent rating interaction

7.4.3.4 Absolute formants of following vowels

LMEMs showed that only F1 and F2 of the following vowel were significantly affected by emphasis (Table 7.34 and Figure 7.62). The two first formant frequencies (F1 and F2) at the onset of vowels in an emphatic environment were significant predictors of emphasis in both FDH-directed speech and ADS as demonstrated by *pot hoc* tests shown in Table 7.35 ($p<0.01$). F1 was found to be higher for /tˤ/ than /t/, whereas F2 was found to be lower for /tˤ/ than /t/. This indicates an approximation of the two first formants in the realization of the emphatic stop sound. However, F3 did not change significantly from /t/ to /tˤ/ as Table 7.35 illustrates ($p>0.05$).

Metric	**Predictor**	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
F1	(Intercept)	6.07	0.08	74.99	70.808	
	Sound /t/	-1.08	0.09	163.71	-11.001	**<0.0001*** **
	Speech FDS	0.07	0.08	167.64	0.93	0.35
	/t/*FDS	0.08	0.12	165.98	0.66	0.509
Random effects	speaker	Variance 0.06	Std.Deviation 0.24			
F2	(Intercept)	9.85	0.09	138.79	108.66	
	Sound /t/	3.51	0.12	170.31	27.62	**<0.0001*** **
	Speech FDS	-0.11	0.108	177.91	-1.04	0.29
	/t/*FDS	-0.109	0.16	174.19	-0.67	0.501
Random effects	speaker	Variance 0.01	Std.Deviation 0.12			
F3	(Intercept)	15.87	0.09	68.33	162.44	
	Sound /t/	-0.06	0.108	163.37	-0.61	0.53
	Speech FDS	-0.102	0.09	166.82	-1.09	0.27
	/t/*FDS	0.04	0.13	165.39	0.35	0.72
Random effects	Speaker	Variance 0.11	Std.Deviation 0.34			

Table 7.34 Results of LMEMs for the effect of speech style on formant frequencies of vowels following stop consonants

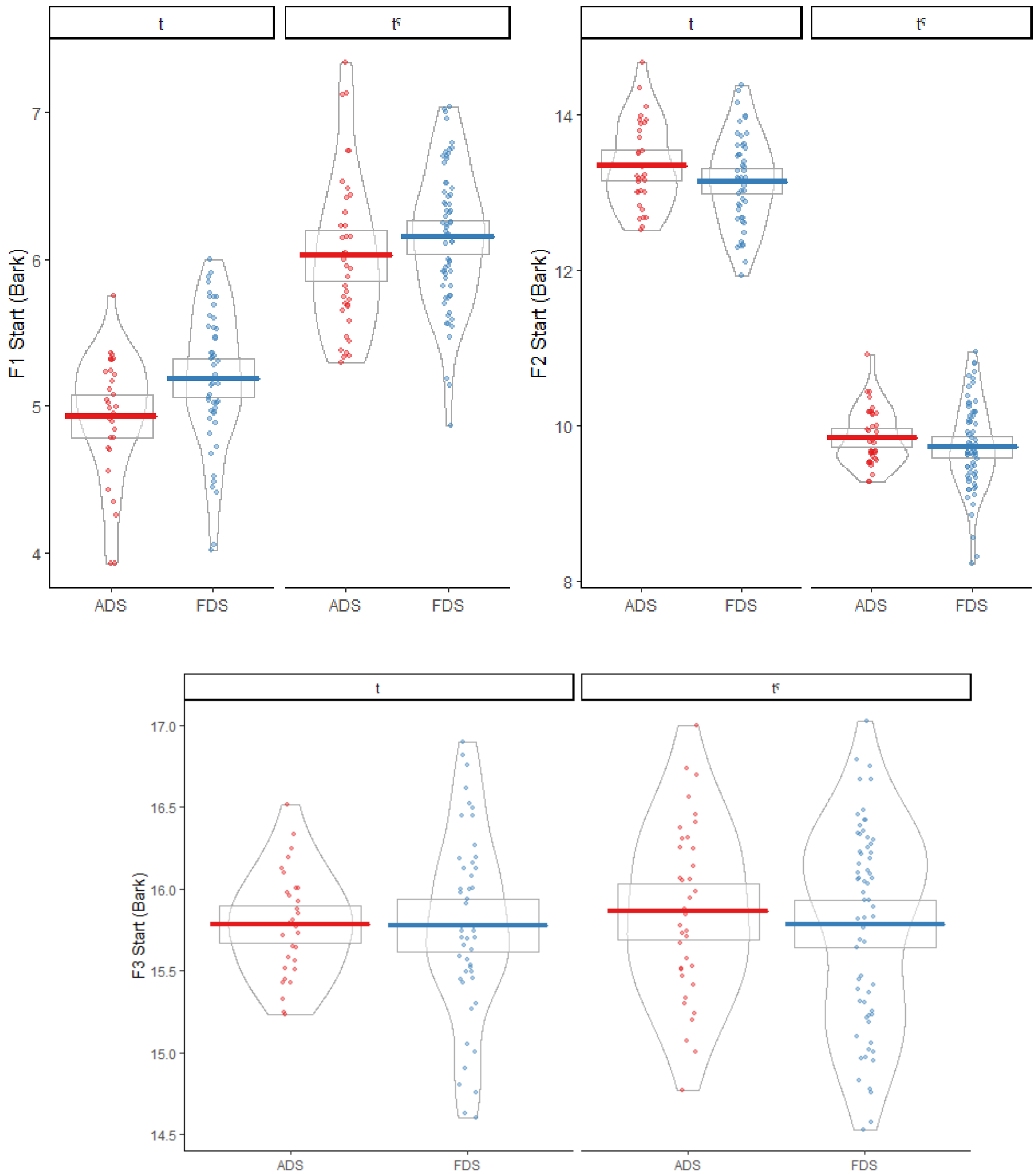


Figure 7.63 Absolute formant frequencies of vowels following /t/ and /tˤ/ within and across speech style

Metric	Sound	Register	emmean	SE	df	lower.CL	upper.CL
F1	$t^{ʕ}$	ADS	6.08	0.08	76.9	5.90	6.25
		FDS	6.16	0.07	46.5	6.00	6.31
	t	ADS	4.99	0.09	90.0	4.81	5.17
		FDS	5.15	0.08	60.6	4.99	5.32
		Contrasts	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	$t^{ʕ}$	ADS-FDS	-0.07	0.08	170	-0.92	0.35
	t	ADS-FDS	-0.16	0.09	171	-1.68	0.09
	ADS	$t^{ʕ}$ -t	1.09	0.100	166	10.88	**<.0001**
	FDS	$t^{ʕ}$ -t	1.00	0.08	173	12.38	**<.0001**
F2	**Sound**	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
	$t^{ʕ}$	ADS	9.85	0.09	137	9.67	10.03
		FDS	9.74	0.07	74	9.59	9.88
	t	ADS	13.36	0.09	158	13.17	13.56
		FDS	13.14	0.08	101	12.98	13.31
		Contrasts	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	$t^{ʕ}$	ADS-FDS	0.11	0.11	181	1.03	0.303
	t	ADS-FDS	0.22	0.12	181	1.804	0.07
	ADS	$t^{ʕ}$ -t	-3.51	0.12	171	-27.29	**<.0001**
	FDS	$t^{ʕ}$ -t	-3.40	0.103	184	-33.09	**<.0001**
	Sound	**Register**	**emmean**	**SE**	**df**	**lower.CL**	**upper.CL**
F3	$t^{ʕ}$	ADS	15.9	0.09	69.8	15.7	16.1
		FDS	15.8	0.08	43.5	15.6	15.9
	t	ADS	15.8	0.104	81.3	15.6	16.0
		FDS	15.8	0.09	56.0	15.6	15.9
		Contrasts	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
	$t^{ʕ}$	ADS-FDS	0.102	0.09	169	1.08	0.28
	t	ADS-FDS	0.05	0.107	170	0.502	0.61
	ADS	$t^{ʕ}$ -t	0.06	0.109	166	0.609	0.54
	FDS	$t^{ʕ}$ -t	0.01	0.08	172	0.19	0.84

Table 7.35 Results of post hoc pairwise comparison tests for the effect of speech style on absolute formant frequencies of vowels following /t/ and /$t^{ʕ}$/

7.4.3.4.1 The effect of foreign accent rating, LoR and religion on stop formant frequencies

LMEMs demonstrated that the effect of foreign accent rating, LoR and religion on absolute formant frequencies of the following vowels is insignificant as shown in Table 7.39 (p>0.05). This indicates the stop contrast is not enhanced in FDH-directed speech by looking at the most robust effects of this contrast, that is F2 of following vowels.

Metric	**Predictor**	**Estimate**	**Std.Error**	**df**	**t-value**	
F1	Intercept	5.14	0.09	32.006	54.35	
	Target Sound:tˤ	1.01	0.07	101.49	13.32	<2e-16 ***
	Accent rating	0.07	0.06	28.77	1.12	0.27
	LoR	0.0018	0.014	27.58	0.12	0.9
	Religion: Non-Muslim	-0.06	0.21	30.75	-0.29	0.77
	Target Sound*Accent rating	-0.06	0.06	100.15	-1.03	0.3
Mixed effects	speaker	Var 0.07	SD 0.28			
F2	Intercept	1.316e+01	1.014e-01	4.354e+01	129.77	
	Target sound:tˤ	-3.417e+00	1.081e-01	1.089e+02	-31.61	<2e-16 ***
	Accent rating	-2.969e-02	7.218e-02	3.867e+01	-0.41	0.68
	LoR	7.026e-03	1.515e-02	2.976e+01	0.46	0.64
	Religion: Non-Muslim	-2.852e-04	2.275e-01	3.461e+01	-0.001	0.99
	Target Sound*Accent rating	9.830e-02	8.745e-02	1.082e+02	1.12	0.26
Mixed effects	speaker	Var 0.03	SD 0.19			

Table 7.36 Results of LMEMs for the effect of LoR on acoustic metrics of /t/ and /tˤ/ in FDH-directed speech

7.4.3.5 Other effects

Beforehand, we found that there was a main effect of speech style on two metrics (kurtosis and intensity) of /tˤ/. To rule out any effects of condition order and phrase boundary on significant changes in these metrics, LMEM results in Table 6.30 show the difference between ADS and FDH-directed speech in condition order 1 (O1) and condition order 2 (O2). LMEMs and Figure 6.54 demonstrate that kurtosis of /$t^{ʕ}$/ is constantly higher in FDH-directed speech than in ADS in both condition orders. This indicates that the difference between FDH-directed speech and ADS in kurtosis of /$t^{ʕ}$/ is not a by-product of who carried out the task first. LMEM revealed that the difference in kurtosis of /$t^{ʕ}$/ was significant only in condition order 2, where the FDH carried out the task second in turn ($p<0.05$). Table 7.36 and Figure 7.63 also show that the difference between FDH-directed speech and ADS with regards to intensity of /$t^{ʕ}$/ is parallel in both condition orders, with FDH-directed speech showing significantly higher intensity in both conditions ($p<0.05$).

Metric	**Order**	**Predictor**	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
/$t^{ʕ}$/							
Kurtosis	O1 FDS/ADS	ADS (baseline)	43.37	19.94	24.33	2.17	
		FDS	33.65	20.44	42.19	1.64	0.107
	O2 ADS/FDS	ADS (baseline)	25.74	20.28	32.34	1.26	
		FDS	45.49	20.65	39.70	2.203	**0.03 ***
Random Effects (O1)	Speaker	Variance 1715	Std.deviation 41.42				
Random Effects (O2)	Speaker	Variance 1216	Std.deviation 34.86				
Intensity	O1 FDS/ADS	ADS (baseline)	49.65	1.36	15.75	36.42	
		FDS	2.802	1.18	38.52	2.36	**0.02 ***
	O2 ADS/FDS	ADS (baseline)	46.69	1.34	31.55	34.66	
		FDS	3.09	1.32	40.07	2.33	**0.02 ***
Random Effects (O1)	Speaker	Variance 11.49	Std.deviation 3.39				
Random Effects (O2)	Speaker	Variance 6.31	Std.deviation 2.51				

Table 7.37 Results of LMEMs for the effect of condition order on kurtosis and intensity of / $t^{ʕ}$/ across register

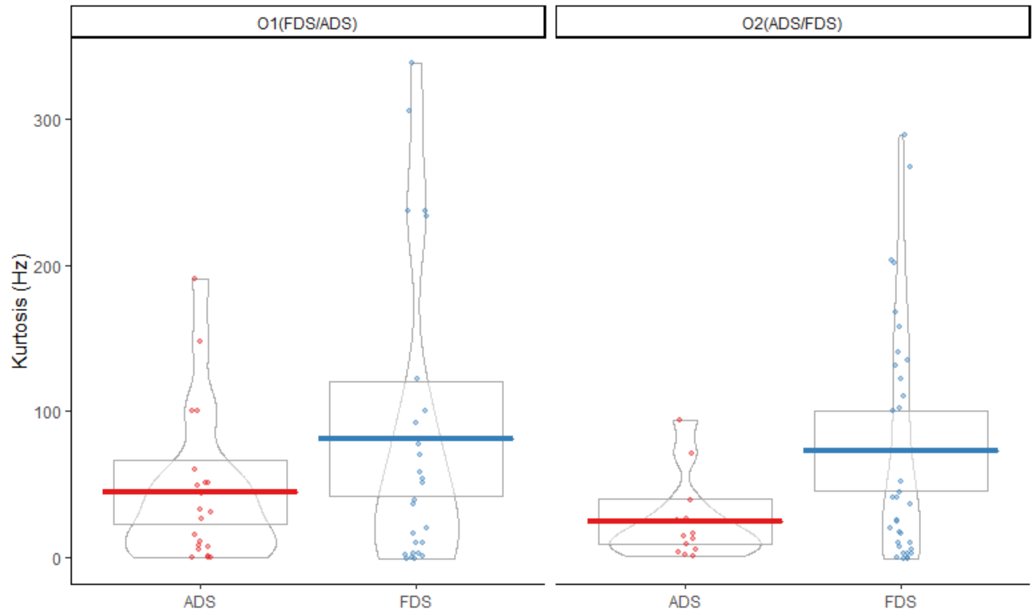


Figure 7.64 Kurtosis values of /tˤ/ within and across speech style in relation to condition order

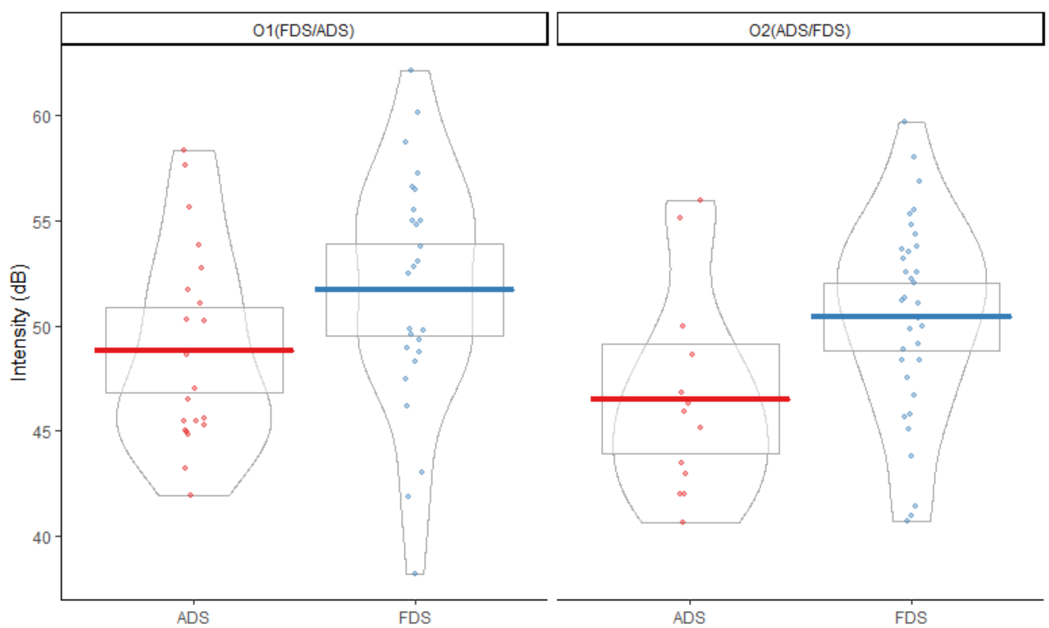


Figure 7.65 Intensity values of /t/ within and between speech style in relation to condition order

Likelihood ratio tests revealed that the full model containing speech style was significantly contributing to the difference between FDH-directed speech and ADS in terms of kurtosis and intensity in /tˤ/ as demonstrated by Table 7.37 ($p<0.05$). This indicates that phrase boundary did not play a role in this difference.

Metric	Model	Df	AIC	BIC	loglik	deviance	Chisq	Df	Pr(>Chisq)
/tˤ/									
Kurtosis	Reduced Model	5	1085.5	1098.2	-537.72	1075.5			
	Full model	6	1080.7	1096	-534.37	1068.7	6.71	1	**0.009 ****
Intensity	Reduced Model	5	566.56	579.28	-278.28	556.56			
	Full Model	6	560.04	575.30	-274.02	548.04	8.52	1	**0.003 ****

Table 7.38 Likelihood ratio tests for the effect of phrase boundary on kurtosis and intensity of /tˤ/

Likelihood ratio tests revealed that the difference between FDH-directed speech and ADS in terms of kurtosis and intensity of /tˤ/ is not a by-product of word stress as the full model containing speech style contributes significantly to this difference as shown in Table 7.38 ($p<0.01$).

Metric	Model	Df	AIC	BIC	loglik	deviance	Chisq	Df	Pr(>Chisq)
/ tˤ /									
Kurtosis	Reduced Model	4	1083	1093.1	-537.47	1075			
	Full Model	5	1078.7	1091.5	-534.36	1068.7	6.21	1	**0.01***
Intensity	Reduced Model	4	566.97	577.14	-279.48	558.97			
	Full Model	5	560.67	573.39	-275.33	550.67	8.29	1	**0.003 ****

Table 7.39 Likelihood ratio tests for the effect of word stress on kurtosis and intensity of /tˤ/

7.4.4 Summary and Discussion

Results from Section 6.4 did not provide support to hypothesis H3 which states that NSs will hyperarticulate their speech acoustically and prosodically to enhance its clarity for their FDH interlocutors. H3a was rejected because results revealed no enhancement or hyperarticulation in the consonantal contrasts /s/ vs. /sˤ/ and /t/ vs. /tˤ/ in FDH-directed speech. H3b was supported to some extent. Intensity of the consonants examined was slightly higher in FDH-directed speech compared to ADS. This proved significant only for /tˤ/. H3c was not supported because consonants were no longer in duration in FDH-directed speech. H4 was not supported either because results revealed no effect of foreign accent rating, LoR or religion on expansion of the fricative and stop contrasts in FDH-directed speech, most of the acoustic metrics examined.

The findings revealed that the only acoustic cues that the NSs might have used to maximize the plain/emphatic sibilant contrast when addressing the FDHs in this study would be F1 and F2 of the following vowel. Results from this section showed that the main acoustic correlates of emphasis for the plain/emphatic sibilants is a raised F1 and a lowered F2 at the onset of the vowel following the emphatic consonant. A raised F1 of vowels is in line with studies discussed in Chapter 4, which found that F1 rising of vowels in an emphatic domain is a consistent exponent of emphasis (e.g Giannini and Pettorino 1982; Zawaydeh 1999; Jongman *et al.* 2011; Jongman *et al.* 2007). F2 lowering of vowels following the emphatic sibilant is in line with the most common finding on this correlate in the literature. Results also revealed that spectral moments, consonant duration and intensity, as well as F3 at the onset of the following vowel, did not play any role in distinguishing the plain/emphatic sibilant contrast. The lack of spectral moments' effect in distinguishing the plain/emphatic continuants is in line with previous research (Bin-Muqbil 2006; Jongman *et al.* 2011). Similarly, as in previous studies, duration was not found to be a significant cue to emphasis in fricatives (Al-Masri and Jongman 2003).

FDH-directed speech and ADS did not differ significantly in center of gravity (mean), standard deviation (variance), kurtosis or F3 of the following vowel for either /s/ or /sˤ/. FDH-directed speech and ADS only exhibited a significant difference in skewness of both /s/ and /sˤ/. This difference indicates that NSs' fricative spectra was more positively skewed when addressing FDHs than it was when addressing the NS. Skewness of both /s/ and /sˤ/ was found to be higher in FDH-directed speech compared to ADS in both condition order 1 and 2, which indicates the lack of effect of this external factor on the present result. With regards to duration, both /s/ and /sˤ/ were shorter in duration in FDH-directed speech compared to ADS, with this difference reaching significance for /sˤ/. FDH-directed speech and ADS did not show any significant difference in intensity for either /s/ or /sˤ/, yet the intensity of both sounds was generally higher in FDH-directed speech than in ADS. Sound × speech style interaction was insignificant for all metrics, indicating the absence of phonetic enhancement in FDH-directed speech.

Contrary to the sibilant plain/emphatic contrast, the first three spectral moments, VOT and intensity of the alveolar stop plain/emphatic contrast appeared to be reliable acoustic cues to emphasis. However, results on the acoustic correlates of emphasis with regard to adjacent vowels were comparable to those of the sibilant plain/emphatic contrast. F1 and F2 but not F3 of adjacent vowels were found to be significant acoustic correlates to emphasis in the /t/ and /tˤ/ contrast. Although evidence for the contribution of spectral moments in classifying plain vs. emphatic consonants has been understudied in the Arabic literature, Jongman *et al.* (2011) report center of gravity (CoG) as the only spectral moment that significantly contributed to emphasis in stop consonants in their study. Similar to results of the current study, they found that CoG of emphatic stop consonants was significantly lower than that of their plain counterparts. Bin Muqbil (2006), on the other hand, found all four spectral moments to significantly contribute to emphasis in stop consonants in his study. His results were in line with results from the present study (except for kurtosis) in which the emphatic stop consonant /tˤ/ was found to have significantly lower CoG, lower standard deviation (SD) and higher skewness. Where F1 of the following vowel was found to significantly predict emphasis, studies reported higher F1 values of vowels following emphatic consonants similar to what the current study found (Giannini and Pettorino 1982; Zawaydeh 1999; Jongman *et al.* 2007; Jongman *et al.* 2011; Al-Tamimi 2017). F2 of vowels following emphatic consonants was found to approximate F1 and thus was considerably lower compared to their plain counterparts. Indeed, F2 of the following vowel has been reported to be a constant acoustic and perceptual correlate of emphasis in the Arabic literature as discussed in chapter 5 (Al-Masri and Jongman 2003; Jongman *et al.* 2007; Al-Tamimi 2017). Also compatible with past research, VOT of the plain stop consonant was found to be significantly longer (aspirated) than that of its emphatic counterpart in both FDH-directed speech and ADS. Interesting results were also revealed with regard to intensity. This metric was found to significantly distinguish emphatic from plain stop contrasts. Additionally, FDH-directed speech exhibited significantly higher intensity for /tˤ/ compared to that of ADS. Due to the absence of research on whether intensity could be considered as an acoustic correlate to emphasis, it is hard to compare results from the current study with previous research. Despite the absence of significant interaction between speech style and sound type, it is interesting to see that intensity of /tˤ/ but not /t/ in FDH-directed speech was exaggerated. This could be attributed to NSs trying to speak louder and with extra vocal effort when producing a sound that is more complex for the FDHs. Taken altogether, results on the robust effect of M1, M2 and M3, VOT and intensity of stop plain/emphatic consonants as well as F1 and F2 of adjacent vowels pointed to the necessity of looking at these metrics in the examination of enhancement (hyperarticulation) of this contrast in FDH-directed speech. However, the current study did not find any significant interaction between speech style and sound type, which suggests the lack of hyperarticulation at the consonantal level.

With regard to the effect of foreign accentedness, LoR and religion, the current study found significant interaction between fricative sound and accent rating with regard to center of gravity and skewness values. The interaction indicated an enhancement of the contrast with regard to these two metrics as /s/ was found to be more fronted and /sˤ/ was found to be more retracted. However, this kind of interaction was not attested with regard to formant frequencies of the following vowels. Thus, it is hard to conclude that there was enhancement of the fricative contrast in FDH-directed speech. Further discussion of these results will be provided in Chapter 8.

The following section will present results on the acoustic examination of vowels in FDH-directed speech and ADS.

7.5 Acoustic Examination of Vowels in FDH-directed Speech

This section presents results regarding the acoustic examination of vowels in FDH-directed speech. Acoustic analysis of Arabic vowels was carried out to test the following hypotheses:

H3. NSs will hyperarticulate their speech acoustically and prosodically to enhance its clarity for their FDH interlocutors.

H3d. Vowel space area will be expanded in FDH-directed speech indicating hyperarticulation.

H3e. Vowel intensity will be higher in FDH-directed speech.

H3f. Vowel duration will be longer in FDH-directed speech.

H3g. $f0$ will have no change in FDH-directed speech.

H4. FDH-directed speech will contain less simplified consonants and consonant clusters or less hyperarticulation as FDHs' LoR increases.

Table 7.40 illustrates the number of vowel tokens within each speech style (FDH-directed speech and ADS). Similar to previous findings in this study, it is interesting to notice that the number of tokens containing the target vowels directed to FDHs was roughly double the number of tokens directed to the NS adult (429 vs. 251).

Vowel	Tokens		Total
	FDH-directed speech	ADS	
/aː/	151	83	234
/iː/	154	87	241
/uː/	124	81	205
Total	429 (63%)	251 (36.9%)	680

Table 7.40 Number of vowel tokens per speech style

7.5.1 Vowel space

An LMEM revealed that the difference in the size of vowel space between the two speech styles is significant ($p<0.01$) (Table 7.41 and Figure 7.65). Vowel space area of FDH-directed speech is larger by 2.67 Barks (0.56, standard error) than that of ADS. This indicates that hyperarticulation of vowel space is evident in FDH-directed speech.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	12.93	0.56	21.84	22.93	
Speech: FDS	2.67	0.12	650.36	20.9	**<2e-16 *****
Mixed effects	speaker	Var 6.46	SD 2.54		

Table 7.41 LMEMs for the effect of speech style on vowel space area

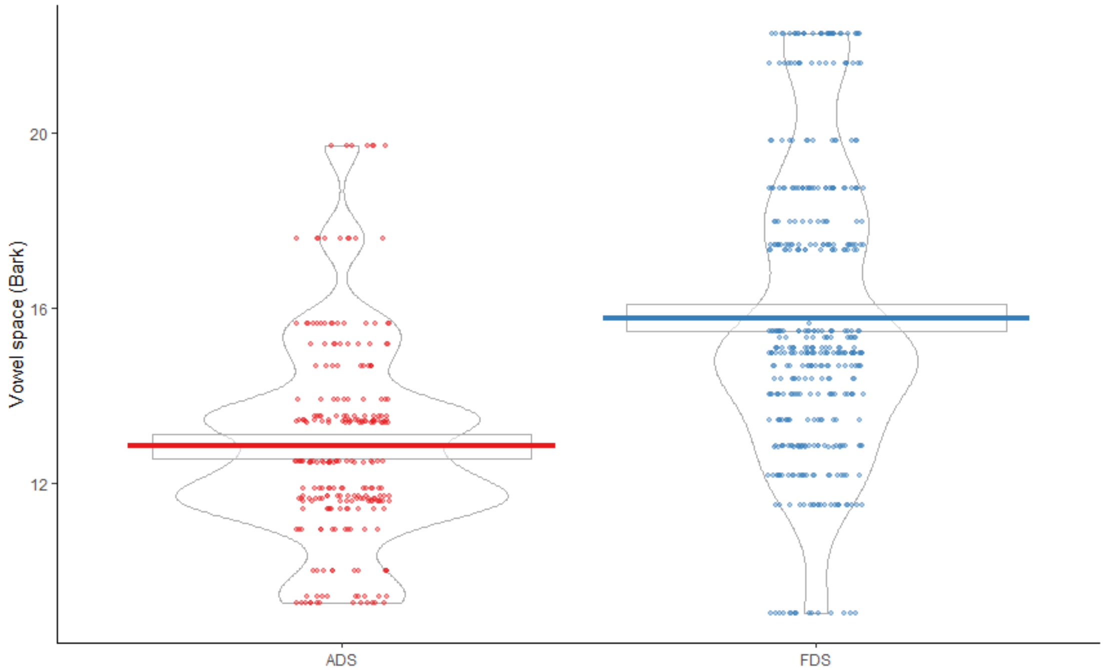


Figure 7.66 The effect of speech syle (FDS vs. ADS) on vowel space

7.5.1.1 The effect of FDHs' foreign accented rating, LoR and religion on vowel space area

An LMEM revealed that vowel space of Omani NSs did not change significantly based on FDH's foreign accentedness score, LoR or religion (Table 7.42). This indicates the absence of effect pertaining to FDHs' linguistic or sociolinguistic characteristics on vowel space of NSs.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	14.38	2.15	21.0006	6.66	
Accent rating	0.2	0.59	21	0.34	0.73
LoR	-0.14	0.14	21.0031	-1.005	0.32
Religion: Non-Muslim	1.23	2.06	21.0053	0.59	0.55
Mixed effects	speaking	Var 9.58	SD 3.09		

Table 7.42 LMEM results for the effect of FDHs' foreign accentedness score, LoR and religion on NSs vowel space

7.5.2 F1

LMEMs demonstrated that F1 for the vowels /iː/, /uː/ and /aː/ in FDH-directed speech was significantly higher than that in ADS as illustrated in Figure 7.66 and Table 7.43 (p<0.01). This provides further evidence for the direction of vowel space expansion along the F1 dimension. These results indicate that all three vowels undergo more jaw lowering in FDH-directed speech compared to ADS. Results of random effects for the three vowels are presented in Table 7.44.

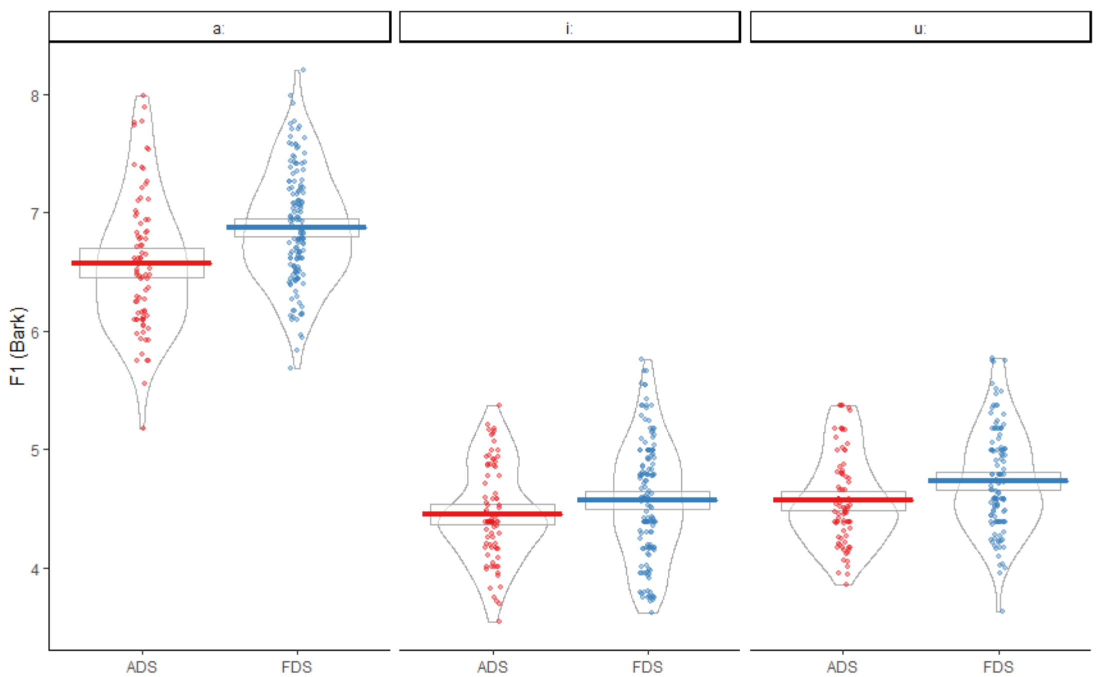


Figure 7.67 F1 averaged means within and across register for the three vowels

	Estimate	Std. Error	df	t value	Pr(>|t|)
/i:/					
ADS (Baseline)	4.43	0.07	26.53	58.17	
FDS	0.11	0.04	221.52	2.605	**0.009 ****
/u:/					
ADS (Baseline)	4.62	0.12	3.1003	38.38	
FDS	0.15	0.04	188.63	3.25	**0.001 ****
/a:/					
ADS (Baseline)	6.59	0.09	24.35	69.99	
FDS	0.205	0.04	213.02	4.58	**<0.0001 *****

Table 7.43 Results of LMEMs for the effect of speech style on F1 of the three vowels

Group	Name	Variance	Std.deviation
/i:/			
Speaker	Intercept	0.09	0.302
Target Word	Intercept	0.001	0.03
/u:/			
Speaker	Intercept	0.03	0.19
Target Word	Intercept	0.037	0.19
/a:/			
Speaker	Intercept	0.14	0.38
Target Word	Intercept	0.002	0.04

Table 7.44 Results of random effects for LMEMs of vowel F1

To rule out the effect of phrase boundary, likelihood ratio tests revealed that the full models (including speech style) were more significant than the reduced models (not including speech style) ($p<0.01$) (Table 7.45). This indicates that speech style contributed significantly to the full models and beyond the influence of phrase boundary. In other words, higher F1 values for the three vowels in FDH-directed speech in this data appear to be the result of speakers modifying their speech sounds based on whether the recipient is a foreigner or a native adult, rather than a mere consequence of other factors.

Model	**Df**	**AIC**	**BIC**	**LogLike**	**Deviance**	**Chisq**	**Df**	**Pr(>Chisq)**
/i:/								
Reduced Model	5	201.72	219.14	-95.85	191.72			
Full Model	6	197.06	217.97	-92.53	185.06	6.65	1	**0.00****
/u:/								
Reduced Model	5	180.48	197.1	-85.24	170.48			
Full Model	6	171.16	191.1	-79.57	159.16	11.32	1	**0.00 *****
/a:/								
Reduced Model	14	222.32	239.60	-106.162	212.32			
Full Model	15	203.74	6829.5	-95.872	191.74	20.58	1	<**0.0001 *****

Table 7.45 Results of likelihood ratio tests for the effect of speech style and phrase boundary on vowel F1

Figures 7.67 and 7.68 show the difference in F1 for the three vowels between FDH-directed speech and ADS based on condition order. It is obvious that F1 for the three vowels in FDH-directed speech is constantly higher than that in ADS regardless of condition order (Table 7.46). However, LMEM demonstrated that the difference between FDH-directed speech and ADS in F1 for all vowels is significant in condition order 1 where the FDH-directed speech was carried out first ($p<0.01$), whereas this difference is only significant for /a:/ in condition order 2 where the FDH carried out the task second ($p<0.01$). The results of random effects are presented in Table 7.47.

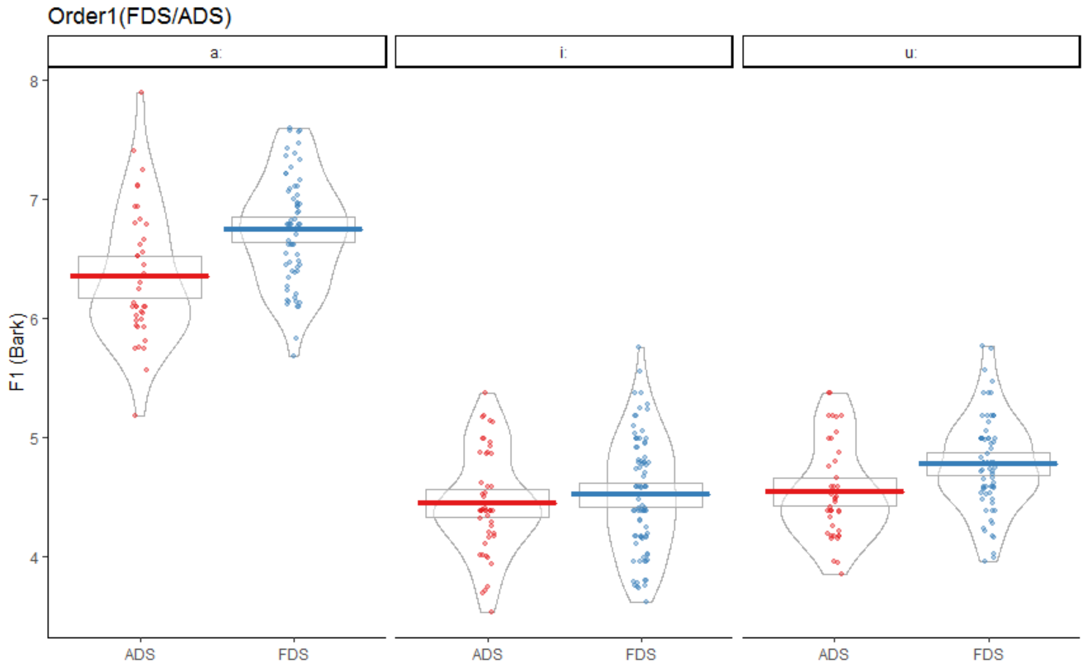


Figure 7.68 F1 for the three vowels across speech registers in relation to task order 1 (FDS/ADS)

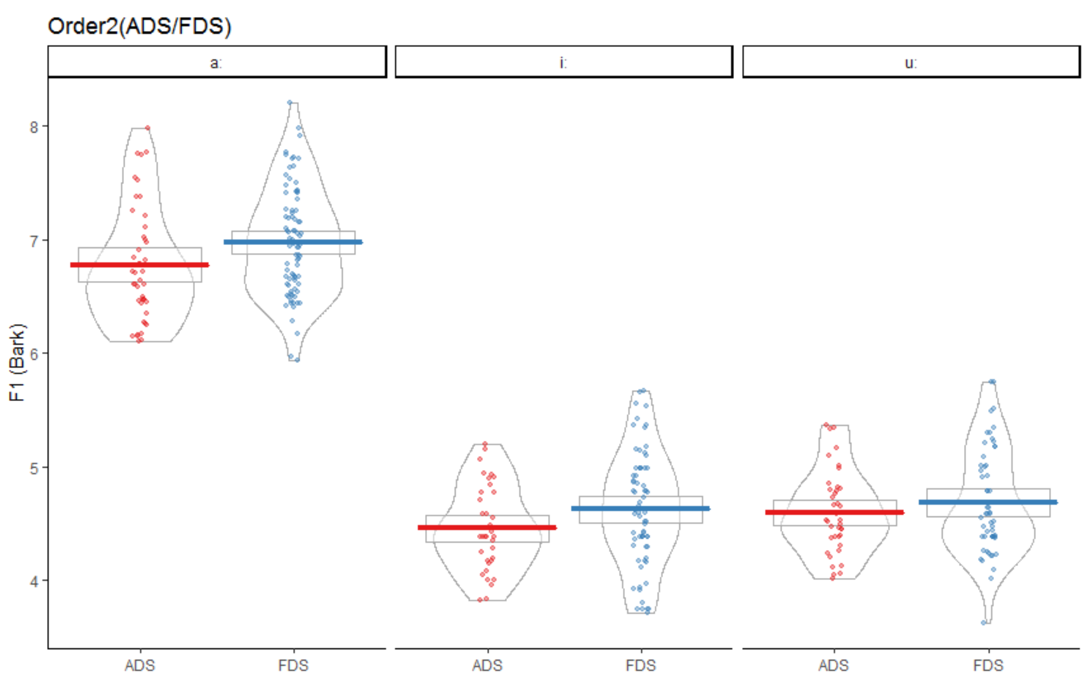


Figure 7.69 F1 for the three vowels across speech style in relation to task order 2 (ADS/FDS)

	Estimate	Std. Error	df	t value	Pr(>|t|)
Task Order1 (FDS/ADS)					
/i:/					
ADS (Baseline)	4.42	0.11	13.802	38.84	
FDS	0.11	0.05	123.58	2.106	0.03 *
/u:/					
ADS (Baseline)	4.61	0.12	4.18	37.18	
FDS	0.23	0.06	99.34	3.62	0.0004 ***
/a:/					
ADS (Baseline)	6.44	0.13	12.74979	48.28	
FDS	0.28	0.06	95.1005	4.34	8.06e-16 ***
Task Order2 (ADS/FDS)					
/i:/					
ADS (Baseline)	4.44	0.102	18.206	43.52	
FDS	0.122	0.07	98.59	1.69	0.09
/u:/					
ADS (Baseline)	4.61	0.13	5.84	35.32	
FDS	0.06	0.06	86.51	0.88	0.37
/a:/					
ADS (Baseline)	6.73	0.11	13.25	56.212	
FDS	0.13	0.05	116.53	2.302	0.0231 *

Table 7.46 Results of LMEMs for the effect of speech style on F1 for the three vowels in condition order 1 and 2

Task Order 1 (FDS/ADS)			
Groups	Names	Variance	Std.Dev.
/i:/			
Speaker	Intercept	0.108	0.32
Target Word	Intercept	0.003	0.06
/u:/			
Speaker	Intercept	0.03	0.17
Target Word	Intercept	0.03	0.17
/a:/			
Speaker	Intercept	0.14	0.38
Target Word	Intercept	0.00	0.06
Task Order 2 (ADS/FDS)			
Groups	Names	Variance	Std.Dev.
/i:/			
Speaker	Intercept	7.385e-02	2.717e-01
Target Word	Intercept	9.488e-19	9.740e-10
/u:/			
Speaker	Intercept	0.04	0.206
Target Word	Intercept	0.03	0.17
/a:/			
Speaker	Intercept	0.13	0.36
Target Word	Intercept	0.00	0.02

Table 7.47 Results of random effects for LMEMs on the effect of condition order on vowel F1

7.5.2.1 The effect of LoR, religion and foreign accent rating on F1

An LMEM revealed that FDHs' foreign accentedness score, LoR and religion did not play any significant role on F1 of the three vowels ($p>0.05$) as shown in Table 7.48. For /i:/, F1 values decrease by 0.05 Barks as foreign accentedness scores rise higher. Also, F1 values decrease by 0.31 Barks when non-Muslim FDHs are the addressees. However, F1 values increase by 0.02 Barks as FDHs' LoR extends. For /u:/, F1 values follow a similar trend to those of /i:/, F1 values decrease bt 0.03 Barks as foreign accent rating increases and by 0.19 Barks when FDHs addressees are non-Muslim. F1 values, on the other hand, increase by 0.02 Barks as FDHs' LoR extends. F1 values of /a:/ increase by 0.01 barks as FDHs' foreign accentedness score rises. They, however drop as FDHs' LoR extends and when FDH addressees are non-Muslims.

Metric	Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
/i:/	Intercept	4.69	0.23	21.19	19.64	
	Accent rating	-0.05	0.06	21.18	-0.81	0.42
	LoR	0.02	0.01	22.31	1.27	0.21
	Religion: Non-Muslim	-0.31	0.23	23.99	-1.35	0.18
Mixed effects	speaker	Var 0.102	SD 0.32			
/u:/	Intercept	4.92	0.19	25.04	25.7	
	Accent rating	-0.03	0.05	23.78	-0.72	0.47
	LoR	0.02	0.01	23.35	1.92	0.06
	Religion: Non-Muslim	-0.19	0.17	21.61	-1.11	0.27
Mixed effects	speaker	Var 0.04	SD 0.21			
/a:/	Intercept	6.83	0.25	20.92	26.73	
	Accent rating	0.01	0.07	20.82	0.15	0.87
	LoR	-0.003	0.01	21.42	-0.2	0.83
	Religion: Non-Muslim	-0.05	0.24	22.14	-0.22	0.82
Mixed effects	speaker	Var 0.12	SD 0.34			

Table 7.48 Results of LMEMs for the effect of LoR, religion, rating and L1 on F1 of the three vowels

7.5.3 F2

LMEMs demonstrated the absence of a significant effect of speech style on F2 of the high front vowel /i:/ and the high back vowel /u:/ as shown in Table 7.49 and Figure 7.69 ($p>0.05$). F2 for /i:/ is lower in FDH-directed speech by 0.08 Barks ± (0.04 standard errors) than that in ADS. Although this difference is negligible, /i:/ in FDH-directed speech appeared to be produced with a slightly retracted tongue compared to ADS. On the other hand, /u:/ in FDH-directed speech appeared to be produced with a slightly advanced tongue compared to ADS as F2 for /u:/ in FDH-directed speech was higher by 0.09 Barks (0.09 standard errors) compared to that in ADS, though this again fell short of significance. Table 7.50 shows random effects results for LMEMs.

Metric	Predictor	Estimate	Std.Error	df	t value	Pr(>|t|)
/i:/	ADS (Baseline)	15.02	0.09	11.33	154.71	
	FDS	-0.08	0.04	218.76	-1.77	0.07
/u:/	ADS (Baseline)	7.81	0.19	4.58	40.46	
	FDS	0.09	0.09	194.26	0.99	0.32

Table 7.49 Results of LMEMs for the effect of speech style on F2 for the vowels /i:/ and /u:/

Groups	Names	Variance	Std.Dev.
/i:/			
Speaker	Intercept	0.09	0.307
Target Word	Intercept	0.01	0.105
Residual		0.109	0.33
/u:/			
Speaker	Intercept	0.03	0.17
Target Word	Intercept	0.09	0.31
Residual		0.41	0.64

Table 7.50 Results of random effects for the effect of speech style of vowel F2

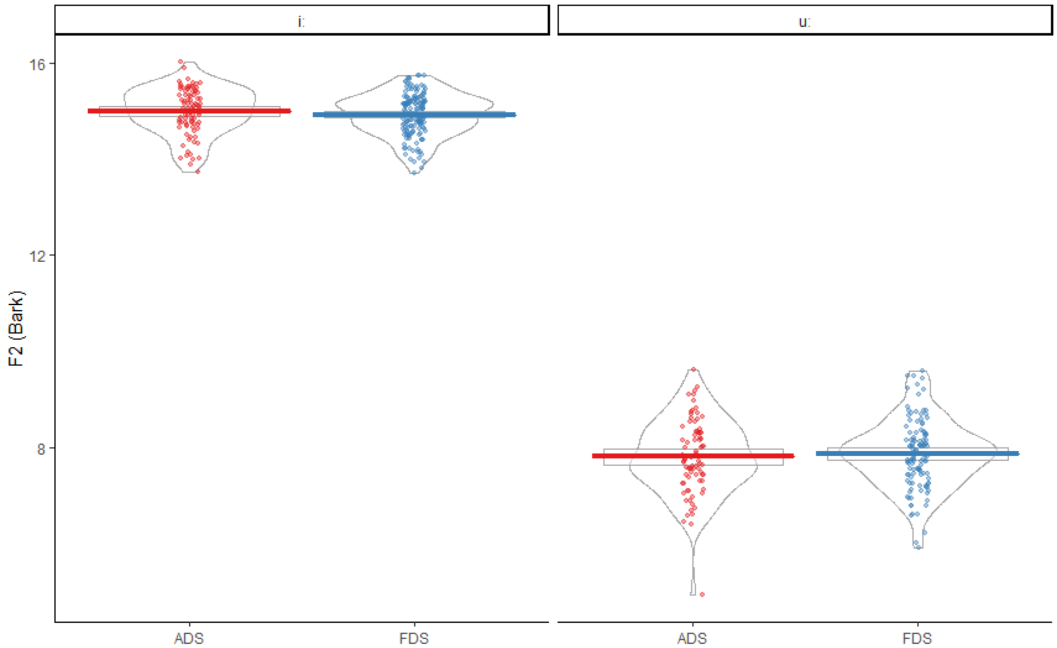


Figure 7.70 F2 for /i:/ and /u:/ across speech styles

Based on F1 and F2 mean values of the three Arabic corner vowels, Figure 7. 70 illustrates the vowel triangle of FDS and ADS. Similar to LMEM results, we can notice more increase in F1 mean values of FDS compared to ADS, which suggest open jaw and stronger vocal effort in production of these vowels in FDS. F2 mean values are similar across speech style as shown in the figure.

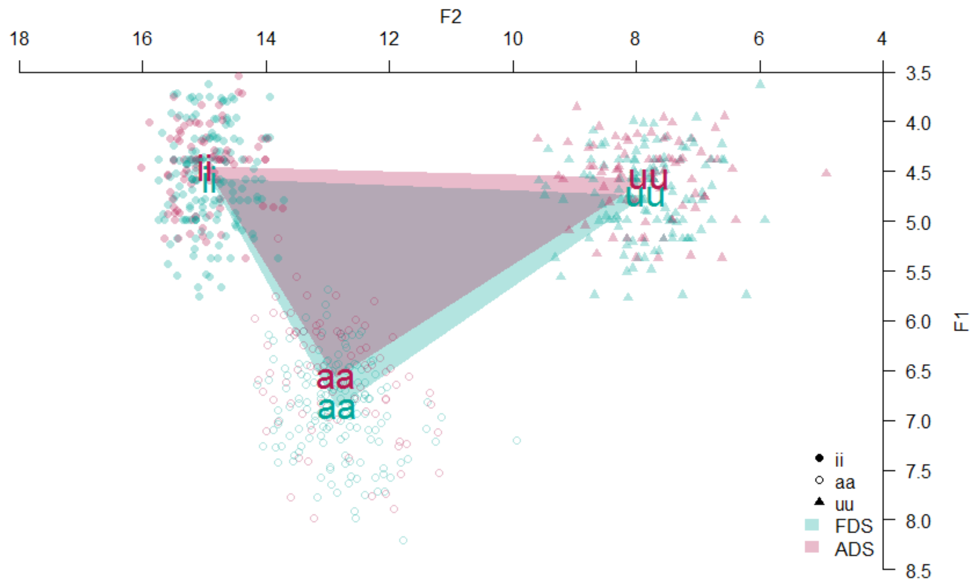


Figure 7.71 Vowel space diagram of Arabic vowels based on F1 and F2 mean values of FDH-directed speech (green shade) and ADS (pink shade)

7.5.3.1 The effect of FDHs' foreign accented rating, LoR, and religion on F2

LMEMs revealed that FDHs' foreign accentedness score, LoR and religion did not play a significant role in F2 variation neither for the front vowel nor for the back vowel in FDH-directed speech as seen in Table 7.51 ($p>0.05$). F2 mean values of /i:/ increase slightly by 0.03 Barks and 0.008 Barks as FDHs' are more native like-score rises and as their LoR extends. However, F2 mean values drop by 0.06 Barks when FDH addressees are non-Muslim. For /u:/, F2 mean values decrease by 0.01 as accent rating score rises. However, F2 mean values increase slightly as FDHs LoR extends and when FDH addressees are non-Muslim.

Metric	Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
/i:/	Intercept	15.09	0.19	20.82	76.48	
	Accent rating	0.03	0.05	20.9	0.68	0.49
	LoR	0.008	0.01	22.51	0.64	0.52
	Religion: Non-Muslim	-0.06	0.19	25.15	-0.33	0.74
Mixed effects	speaker	Var 0.064	SD 0.25			
/u:/	Intercept	7.98	0.25	28.36	31.59	
	Accent rating	-0.01	0.06	23.1	-0.19	0.84
	LoR	0.01	0.01	24.68	1.18	0.24
	Religion: Non-Muslim	0.04	0.22	21.18	0.18	0.85
Mixed effects	speaker	Var 0.02	SD 0.17			

Table 7.51 Results of LMEMs for the effect of LoR on F2 in FDH-directed speech

7.5.4 Intensity

LMEM revealed that intensity of vowels in FDH-directed speech (Mean: 64.21 dB) was significantly higher than that in ADS (Mean: 62.65 dB) as can be seen in Table 7.52 and Figure 7.71 ($p<0.01$). This suggests that speech directed to FDHs was louder than that directed to the NS adult. Table 7.53 presents random effects results.

Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
ADS (Baseline)	62.65	0.85	27.39	73.54	
FDS	1.66	0.28	653.27	5.73	**1.46e-08 *****

Table 7.52 Results of LMEM for the effect of speech style on vowel intensity

Groups	Name	Variance	Std.Dev
Speaker	Intercept	7.77	2.78
Target Word	Intercept	2.95	1.72

Table 7.53 Results of random effects of LMEM for the effect of speech style on intensity

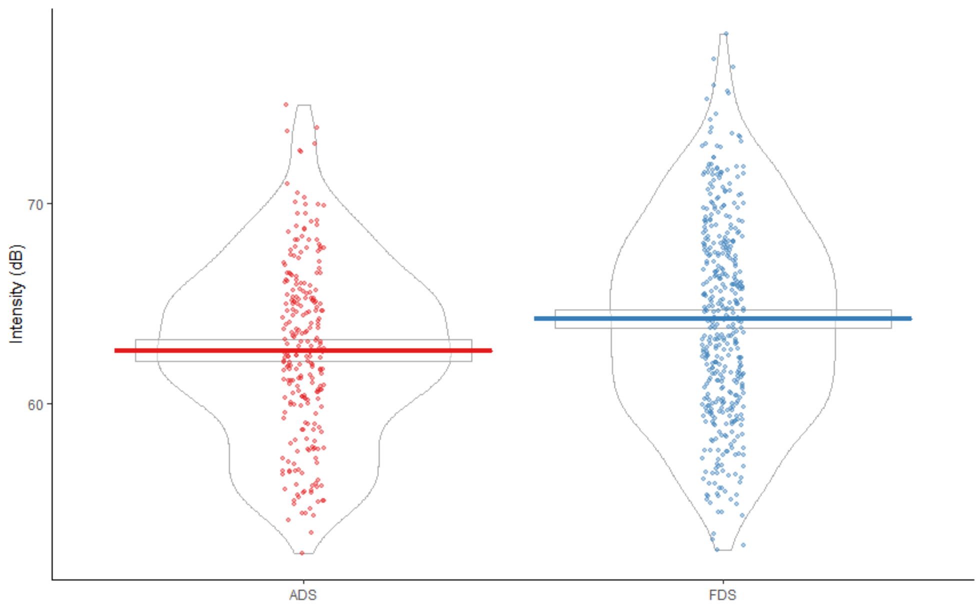


Figure 7.72 Vowel intensity across speech style

To rule out any influence of phrase boundary, a likelihood ratio test revealed that the full model with the effect in question (speech style) was more significant than the reduced model ($p<0.01$) (Table 7.54). This indicates that higher intensity rates in FDH-directed speech were a consequence of speakers changing their speech based on whether the addressee was a foreigner or a NS adult and not a mere consequence of whether the vowel was at a phrase boundary or not.

Model	**Df**	**AIC**	**BIC**	**LogLike**	**Deviance**	**Chisq**	**Df**	**Pr(>Chisq)**
Reduced Model	5	3778	3800.6	-1884	3768			
Full Model	6	3746.5	3773.6	-1867.2	3734.5	33.55	1	**6.929e-09 *****

Table 7.54 Results of likelihood ratio test for the effect of speech style and prosodic boundary on vowel intensity

Figure 7.72 demonstrated that vowel intensity of FDH-directed speech was higher than that of ADS regardless of whether the NS addressed the FDH first or second. LMEM revealed that vowel intensity in FDH-directed speech was significantly higher than that of ADS (p<0.01) in both condition order 1 and 2 (Table 7.55). This demonstrates that higher vowel intensity in FDH-directed speech was not a by-product of who the NS addressed first in the task, but rather was a consequence of who the interlocutor was. Table 7.56 shows results of random effects.

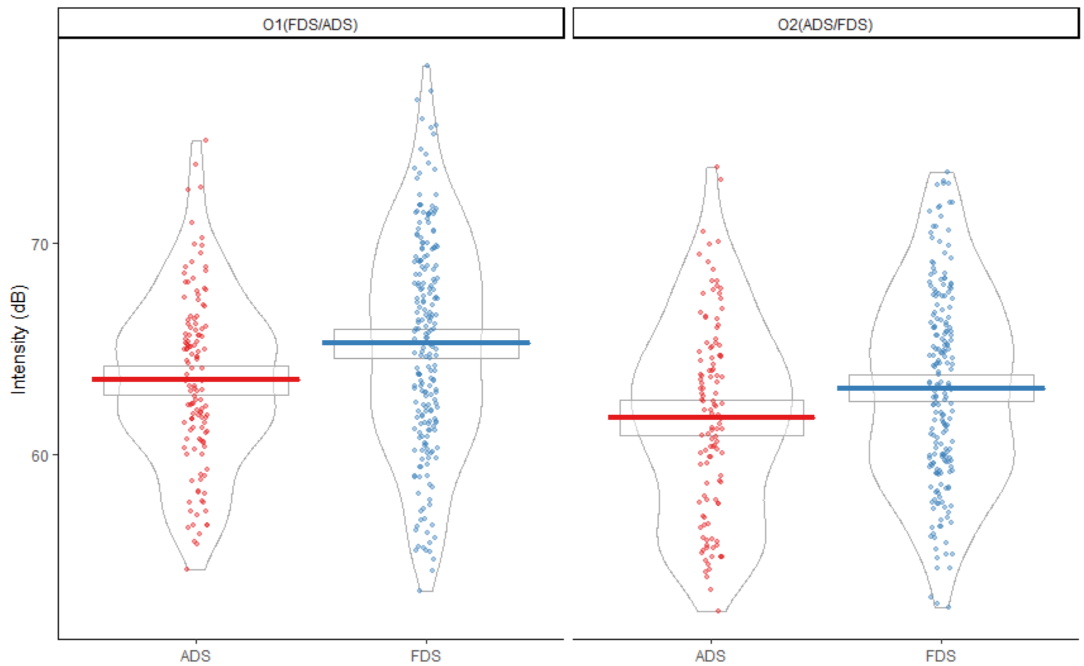


Figure 7.73 The effect of speech style on intensity in task order1 and task order 2

	Estimate	Std.Error	df	t value	Pr(>|t|)
Task Order1 (FDS/ADS)					
ADS (Baseline)	63.64	1.05	19.27	60.44	
FDS	1.79	0.402	334.67	4.45	**1.14e-05 *** **
Task Order2 (ADS/FDS)					
ADS (Baseline)	61.803	0.99	19.706	61.92	
FDS	1.51	0.41	312.16	3.65	**0.0003 *** **

Table 7.55 LMEMs results for the effect of speech syle on vowel intensity in condition orders 1 and 2

Groups	Names	Variance	Std.Dev.
	Task Order 1 (FDS/ADS)		
Speaker	Intercept	7.45	2.73
Target Word	Intercept	3.01	1.73
	Task Order 2 (ADS/FDS)		
Speaker	Intercept	6.23	2.49
Target Word	Intercept	2.91	1.708

Table 7.56 Random effects results of LMEMs for the effect of speech syle on vowel intensity in condition orders 1 and 2

7.5.4.1 The effect of FDHs' foreign accent rating, LoR and religion on intensity

An LMEM revealed that the effect of FDHs' foreign accentedness score, LoR and religion on vowel intensity was insignificant as shown in Table 7.57 ($p>0.05$). Mean intensity value drops by 0.13 dB when FDHs are more native-like and by 2.19 dB when FDH addressees are non-Muslim. However, mean intensity value increases as FDHs' LoR extends.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	65.81	2.12	21.01	31.02	
Accent rating	-0.13	0.58	20.9	-0.22	0.82
LoR	0.12	0.14	21.39	0.91	0.37
Religion: Non-Muslim	-2.19	2.04	21.79	-1.07	0.29
Mixed effects	Speaker	Var 8.43	SD 2.9		

Table 7.57 Results of LMEM for the effect of foreign accent rating, LoR and religion on vowel intensity in FDH-directed speech

7.5.5 Duration

LMEM revealed that vowels in FDS were significantly shorter in duration (mean: 135 ms) than vowels in ADS (mean: 145.6 ms) as shown in Table 7.58 and Figure 7.73 ($p<0.01$). Vowels in FDH-directed speech were shorter by 9.41 ms ± 2.82 (standard errors) than those in ADS. Results of random effects are presented in Table 7.59.

Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
ADS (Baseline)	145.57	5.63	19.45	25.83	
FDS	-9.41	2.82	659.089	-3.33	**0.0009 *****

Table 7.58 Results of LMEM for the effect of speech style on vowel duration

Groups	Name	Variance	Std.Dev
Speaker	Intercept	174	13.19
Target Word	Intercept	171.5	13.10

Table 7.59 Results of random effects of LMEM for the effect of speech style on vowel duration

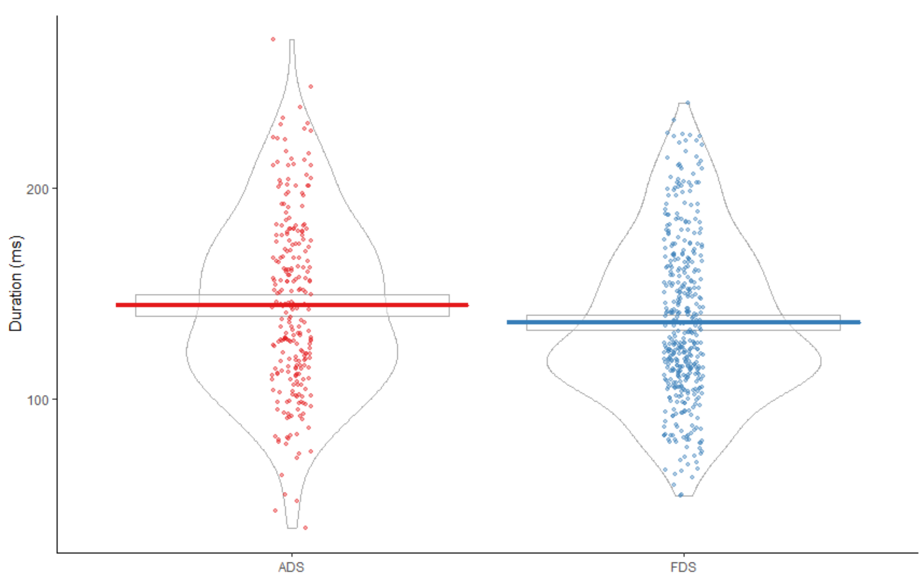


Figure 7.74 Vowel duration across speech registers

A likelihood ratio test revealed that the full model was significant (p<0.01), indicating that speech style was primarily causing the difference in duration between FDH-directed speech and ADS and not phrase boundary (Table 7.60).

Model	**Df**	**AIC**	**BIC**	**LogLike**	**Deviance**	**Chisq**	**Df**	**Pr(>Chisq)**
Reduced Model	5	6812	6834.6	-3401	6802			
Full Model	6	6800.6	6827.8	-3394.3	6788.6	13.33	1	**0.0002*** **

Table 7.60 Likelihood ratio test results for the effect of phrase boundary on duration

LMEMs revealed that vowels in FDS were significantly shorter in duration than those in ADS in both task orders 1 and 2, as shown in Figure 7.74 and Table 7.61 (p>0.05). Accordingly, task order does not appear to cause any effect on how NSs modified their speech based on who they addressed first. Results of random effects are displayed in Table 7.62.

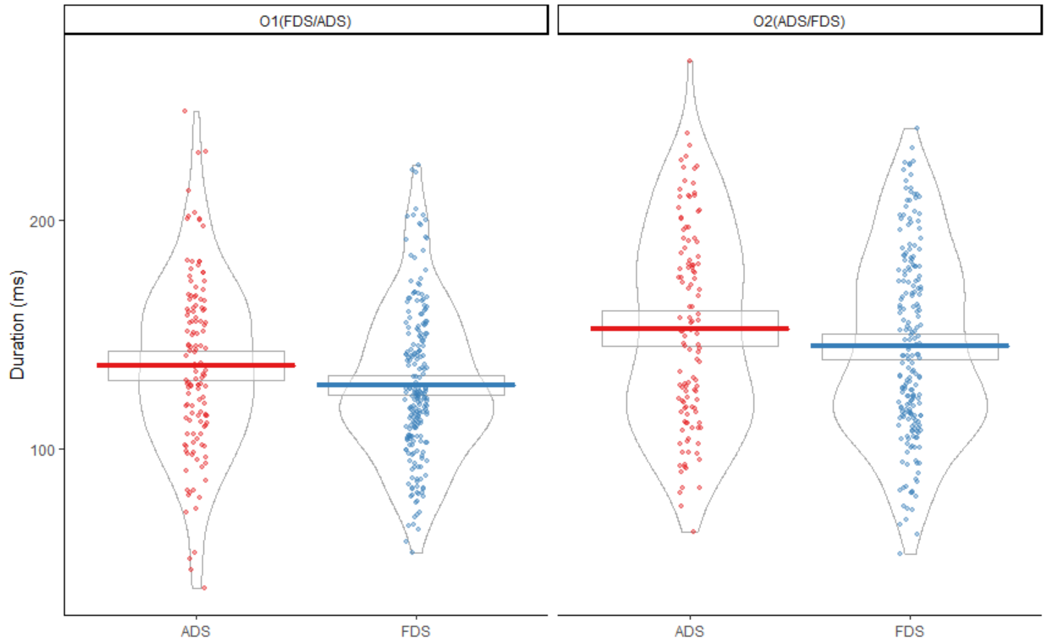


Figure 7.75 Vowel duration across speech style in task orders 1 and 2

Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
Task Order 1 (FDS/ADS)					
ADS (Baseline)	138.809	5.94	20.19	23.34	
FDS	-8.43	3.5	336.29	-2.37	**0.01 ***
Task Order 2 (ADS/FDS)					
ADS (Baseline)	153.21	7.32	15.502	20.92	
FDS	-9.31	4.38	317.29	-2.12	**0.03 ***

Table 7.61 Results of LMEMs for the effect of speech register on vowel intensity in task orders 1 and 2

Groups	Names	Variance	Std.Dev.
	Task Order 1 (FDS/ADS)		
Speaker	Intercept	180.4	13.43
Target Word	Intercept	100.1	10.01
Residual		1004.5	31.69
	Task Order 2 (ADS/FDS)		
Speaker	Intercept	85.13	9.22
Target Word	Intercept	305.40	17.47
Residual		1404.65	37.47

Table 7.62 Results of random effects of LMEMs for the effect of speech register on vowel intensity in task orders 1 and 2

Since vowels in FDH-directed speech were significantly shorter than vowels in ADS, this raised the question as to what could trigger such vowel shortening in FDH-directed especially since vowel duration has been found to undergo significant lengthening in similar speech registers including clear speech (e.g. Pitcheny *et al.* 1986; Ferguson and Kewley-Port 2002), IDS (e.g. Cristia and Seidl 2014) and FDS (e.g. Scarborough *et al.* 2007; Kangatharan 2014). With other factors such as phrase boundary and condition order showing no contribution to this reduction, word repetition was considered as a potential source for this result. A host of observations in the literature have pointed to a tendency by talkers to shorten content words that they have previously produced in their conversations (e.g Folwer and Housum 1987; Fowler 1988; Anderson and Howarth 2002; Rodriguez-Cuadrado *et al.* 2018). This is because second mention words become highly predictable from the context, and thus are more likely to be phonetically reduced (Anderson and Howarth 2002). In the present study, there were more word repetitions in FDH-directed speech than in ADS (see Table 6.34), which could have contributed to this reduction in vowel duration in FDH-directed speech, provided that the task used to elicit data lent itself to spontaneous interactions. However, one might argue that based on the H&H theory, speakers constantly evaluate their listeners' needs for clear signal details and adapt their speech a long a continuum of hyper- and hypo-speech depending on communicative constraints, and thus contrastive sounds might be maximized when talkers repeat utterances that contain sounds that have been mistaken for similar sounds (e.g. Maniwa *et al.*2009). However, Ohala (1994) did not find evidence supporting this assumption in his study. There was no difference in vowel duration and other acoustics as a function of feedback (confusion or misperception) from the listener. Thus, to find out whether the observed difference in vowel duration was a by-product of word repetition, all word repetitions were excluded from both FDH-directed speech and ADS and only first mention target words were submitted to a LMEM. The results indicated that the difference in vowel duration between FDH-directed speech and ADS was insignificant ($p>0.05$) as exhibited in Table 7.63 and Figure 7.75. Vowels in FDH-directed speech were slightly shorter than those of ADS but only by 3.1 ms ± 3.43 (standard errors). Results of random effects are shown in Table 7.64.

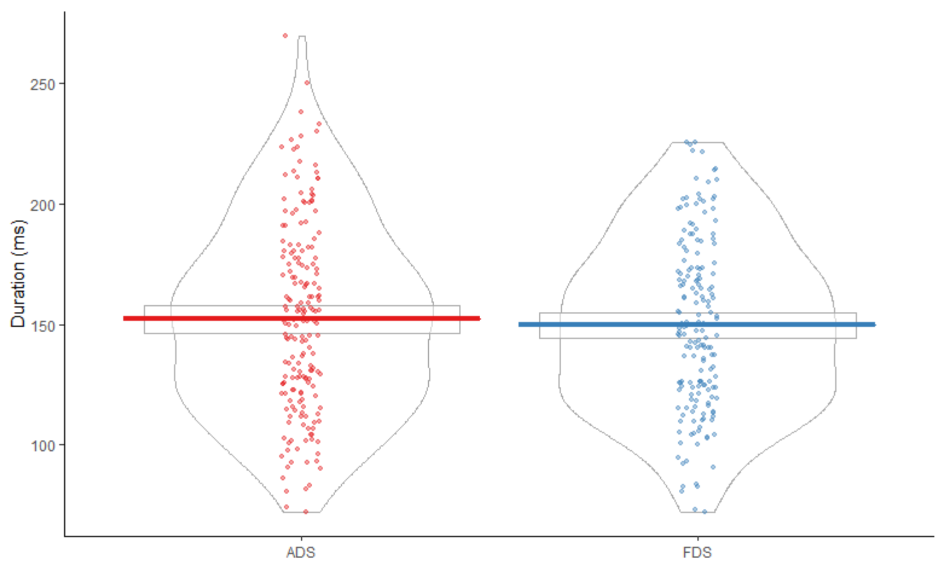


Figure 7.76 Vowel duration of FDH-directed speech and ADS after excluding word repetitions

Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
ADS (Baseline)	151.51	6.07	24.94	77.96	
FDS	-3.105	3.43	-0.905	-0.06	0.36

Table 7.63 Results of LMEM for the effect of speech style on vowel duration after excluding word repetitions

Groups	Name	Variance	Std.Dev
Speaker	Intercept	118.9	10.90
Target Word	Intercept	236.8	15.39

Table 7.64 Results of random effects of LMEM for the effect of speech style on vowel duration after excluding word repetitions

7.5.5.1 The effect of FDHs' foreign accent rating LoR and religion on duration

FDHs' foreign accentedness score, LoR and religion did not exert any significant effects on vowel duration as confirmed by LMEM results in Table 7.65 ($p>0.05$). Vowel duration in FDS drops by 2.84 ms when FDHs' sound more native-like. It also drops by 0.004 ms as FDHs' LoR extends. When FDH addressees are non-Muslim, vowel duration in FDS drops by 8.46 ms.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	132.69	11.1	21.54	11.92	
Accent rating	-2.84	3.07	21.22	-0.92	0.36
LoR	-0.004	0.74	22.54	-0.005	0.99
Religion: Non-Muslim	-8.46	10.89	23.85	-0.77	0.44
Mixed effects	Speaker	Var 191.3	SD 13.83		

Table 7.65 Results of LMEM for the effect of foreign accent rating, LoR and religion on vowel duration in FDS

7.5.5 Fundamental frequency (f0)

LMEM showed that *f*0 of vowels in FDH-directed speech was significantly higher than that in ADS as illustrated in Table 7.66 and Figure 7.76 below ($p<0.01$). Table 7.67 shows results of random effects.

Predictor	**Estimate**	**Std. Error**	**df**	**t-value**	**Pr(>|t|)**
ADS (Baseline)	221.72	4.34	32.53	51.06	
FDS	16.42	2.22	657.42	7.37	**5.05e-13 *****

Table 7.66 Results of LMEM for the effect of speech style on F0

Groups	**Name**	**Variance**	**Std.Dev**
Speaker	Intercept	272.43	16.51
Target Word	Intercept	30.36	5.51
Residual		758.75	27.55

Table 7.67 Results of random effects of LMEM for the effect of speech style on *f*0

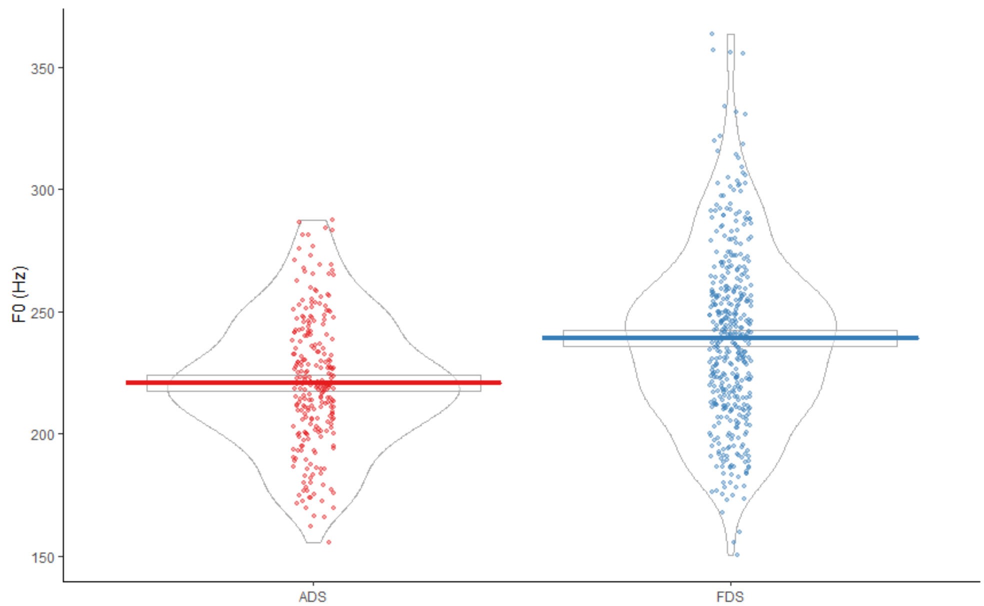


Figure 7.77 *f*0 across speech style

A likelihood ratio test demonstrated that a model including speech style was more significant than a model without this effect as Table 7.68 exhibits ($p<0.01$), indicating that the observed difference in *f*0 between FDH-directed speech and ADS was primarily caused by who the NS was addressing rather than a mere outcome of prosodic boundary.

Model	Df	AIC	BIC	LogLike	Deviance	Chisq	Df	Pr(>Chisq)
Reduced Model	5	6566.6	6589.3	-3278.3	6556.6			
Full Model	6	6516.2	6543.3	-3252.1	6504.2	52.443	1	**4.43e-13 *****

Table 7.68 Results of likelihood ratio tests for the effect of prosodic boundary on f0

Figure 7.77 and Table 6.69 demonstrate that *f*0 of vowels in FDS is significantly higher than that in ADS in both condition orders 1 and 2 ($p<0.01$). Accordingly, condition order does not appear to play a role in the difference in vowel *f*0 between FDS and ADS. Results of random effects are provided in Table 7.70.

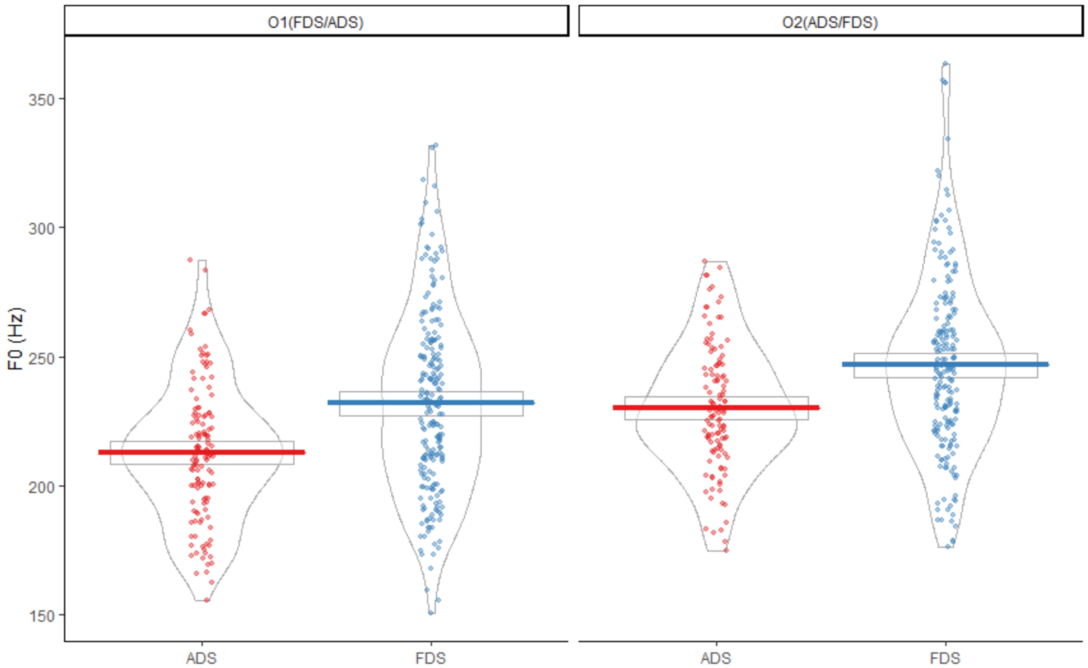


Figure 7.78 Mean *f*0 values across speech style in condition orders 1 and 2

Predictor	Estimate	Std.Error	df	t-value	Pr(>|t|)
Task Order 1 (FDS/ADS)					
ADS (Baseline)	213.19	5.78	16.99	36.85	
FDS	17.503	3.01	337.75	5.81	**1.43e-08 *****
Task Order 2 (ADS/FDS)					
ADS (Baseline)	230.32	4.99	19.83	46.14	
FDS	15.76	3.28	316.98	4.79	**2.5e-06 *****

Table 7.69 Results of LMEMs for the effect of speech style on f0 in task orders 1 and 2

Groups	Names	Variance	Std.Dev.
Task Order 1 (FDS/ADS)			
Speaker	Intercept	261.39	16.16
Target Word	Intercept	35.96	5.99
Task Order 2 (ADS/FDS)			
Speaker	Intercept	155.83	12.48
Target Word	Intercept	36.34	6.02

Table 7.70 Results of random effects of LMEMs for the effect of speech style on f0 in task orders 1 and 2

7.5.5.1 The effect of FDHs' foreign accented rating, LoR and religion on f0

Table 7.71 shows that FDHs' foreign accented rating, LoR and religion do not play any significant role in *f*0 variation in FDH-directed speech (p>0.05). Vowel *f*0 of FDH-directed speech remained roughly constant regardless of FDHs' heterogenous nature. F0 mean value increased by 3.71 Hz as FDHs' foreign accentedness rating increased. It similarly increased as FDHs' LoR extended. When FDHs' are non-Muslim, *f0* mean values increased by 7.57 compared to when they were Muslims.

Predictor	**Estimate**	**Std.Error**	**df**	**t-value**	**Pr(>|t|)**
Intercept	246.54	13.18	20.89	18.7	
Accent rating	3.71	3.64	20.72	1.01	0.32
LoR	0.18	0.87	21.46	0.21	0.83
Religion: Non-Muslim	7.57	12.76	22.1	0.59	0.55
Mixed effects	Speaker	Var 308.7	SD 17.57		

Table 7.71 LMEM results for the effect of FDHs' foreign accentedness score, LoR and religion on vowel *f*0

7.5.6 Summary and Discussion

Results from the previous sections confirmed H3 which states that NSs will hyperarticulate their speech acoustically and prosodically to enhance its clarity for their FDH interlocutors. However, results on the acoustic properties of vowels did not provide support to H4 which predicted that hyperarticulation will be more evident in speech directed to FDHs who are more foreign accented, who are non-Muslims and whose LoR is shorter.

Results on vowel space confirmed H3d which predicted that vowel space will be hyperarticulated in FDH-directed speech. Results showed that vowel space of FDS was significantly larger than that of ADS, confirming with previous studies on FDS (Scarborough *et al.* 2007; Uther *et al.* 2007), clear speech (Bradlow *et al.* 1996; Ferguson and Kewley-Port 2002) and IDS (khul *et al.* 1997; Liu *et a.l* 2003). Mean vowel F1 values in FDH-directed speech also proved to be significantly higher than those of ADSs. Higher F1 mean for vowels reflect more opening of the jaw and an increase in vocal effort (Lienard and DiBenedetto 1999). In contrast with F1, F2 for the front and back vowels remained statistically identical in FDH-directed speech and ADS. In fact, the vowel space of FDH-directed speech was narrower along the F2 dimension compared to ADS, although this trend did not reach significance. The rise in F1 for all vowel categories in this study was also reported in a study by Ferguson and Kewley-Port (2002) for clear speech. Furthermore, the general trend of vowel space expansion in FDH-directed speech reflected by the change in F1 only is in line with that reported by Dodane and Al-Tamimi (2007) for IDS. To this end, results on vowel space expansion in this study support the hyperarticulation hypothesis.

With regard to vowel intensity, results confirmed H3e which predicted that vowels in FDH-directed speech will be higher in intensity (louder). This demonstrated a significant effect of speech style on loudness, triggered by an increase in vowel intensity of FDH-directed speech. Results have also indicated that the prosodic boundary and task order did not play any significant role in varying vowel intensity across speech styles. Results for the effect of LoR on changes in vowel properties showed that LoR did not play any significant role in modifications of F1, F2, duration, $f0$ or intensity. In other words, native speakers did not modify their speech based on their FDHs' Arabic language experience but rather maintained the same acoustic properties regardless of their FDHs Arabic competence.

Results of vowel duration did not confirm H3f which predicted vowel lengthening in FDH-directed speech. Results showed an absence of a significant difference in vowel duration between FDH-directed speech and ADS. In fact, vowels of FDH-directed speech were found to be shorter in duration compared to those of ADS. While this result was in line with that reported in most studies of FDS (Knoll and Scharrer 2007; Uther *et al.* 2007; Knoll *et al.* 2015), it did not coincide with those reported in other FDS studies (Biersack *et al*. 2005; Scarborough *et al.* 2007; Kangatharan 2014), IDS (Fernald and Simon 1984; Andruski and Khul 1996) or clear speech (Picheny *et al.* 1986; Ferguson and Kewley-Port 2002). To explain the reduction of vowel duration in the current study, I considered word repetitions in FDH-directed speech. Data analyses from this section showed that FDH-directed speech had more repetitions indicated by a larger number of target word-tokens compared to ADS. This is in line with previous studies that have revealed that foreigner talk is characterized by more repetitions and clarifications (Ferguson 1971; Taron 1980; Henzl 1979; Arthur *et al.* 1980). There are two contradictory arguments that might interfere in this regard though. The first is the well-established evidence for word shortening as a result of repetition in spontaneous speech (e.g. Fowler and Housum 1987; Fowler 1988; Bell *et al.* 2002; Baker and Bradlow 2009) and the other is the assumption that a listener's feedback (e.g. misperception or confusion of sounds) could trigger clearer enunciations of speech sounds in repeated utterances (Lindblom 1990; Maniwa *et al.* 2009). Unfortunately, because the current study was not designed to test the second argument, it was hard to make further analyses with the data to check for the validity of such an assumption. Furthermore, evidence exists as to the absence of a change in vowel duration of clear speech as a function of the listener's feedback (Ohala 1994). Since the method adopted in the current study to collect data was semi natural in which speakers had the chance to repeat words for further negotiation and there is well established evidence for reduction of content words as a function of repetitions, it was well justified to test the first argument. Results revealed no difference in vowel duration between FDH-directed speech and ADS after excluding repeated content words. Results revealed that the difference in vowel duration between ADS and FDH-directed speech dropped from 9.4 ms to 3.1 ms when repeated words were removed from the analysis. The difference in vowel duration between the two speech registers also truned to be insignificant. This indicates that repetitions seemed to play a role in reducing vowel duration.

With regard to vowel *f*0, results did not support H3g which predicted that *f*0 will remain unchanged in both speech styles. Results revealed a significant effect of speech style on pitch triggered by an increase in *f*0 of vowels in FDH-directed speech. Additionally, this finding does not coincide with findings reported in FDS studies (Biersack *et al.* 2005; Uther *et al.* 2007; Kangatharan 2014) in which *f*0 has been found to show no change across speech styles. In fact, exaggerated mean *f*0 is a robust characteristic of IDS and clear speech, and hence higher *f*0 mean of FDH-directed speech in this study conforms to findings of IDS studies (Fernald and Simon 1984; Papousek *et al.* 1987; Kuhl *et al*. 1997; Uther *et al.* 2007) and clear speech (Uchanski 2004). Additionally, results have shown that the robust effect of speech style on *f*0 is independent of other factors such as prosodic boundary and task order. Further discussion on these results will be presented in Chapter 9.

Chapter 8. Learners' Perception and Production of Arabic Speech

8.0 Introduction

This chapter will present results regarding discrimination of consonantal contrasts and the production of consonants and consonant clusters by foreign domestic helpers (FDHs). It will also provide results with regard to the social and cognitive factors (LoR, L1 formal education and L2 literacy) that influence FDHs' speech learning. Section 8.1 will present results regarding FDHs' accuracy and d prime scores in the AX-forced choice task and the factors that play a role in their performance. Section 8.2 will provide results concerning FDHs' production of Arabic consonants and the factors that affect their performance. Section 8.3 will provide results regarding FDHs production of onset and coda clusters and the factors that affect their performance.

8.1 FDHs' Discriminability of Consonantal Contrasts

This section provides results on FDH's discriminability of consonant contrasts and the factors that play a role in their performance. This examination was set to test the following hypotheses:

> **H5.** FDHs' perceptual sensitivity towards Arabic phonemic contrasts will not reach native-like performance.
>
> **H6.** Foreigners with longer LoR, and who have more formal schooling and who are literate in their L2 will be more sensitive to Arabic phonemic contrasts.

8.1.1 Descriptive Statistics

Table 8.1 and figure 8.1 show the distribution of scores FDHs and NSs obtained in the AX forced-choice task. We can notice that, generally, FDHs' discriminability of the *different* pairs was not low. The percentage of their *Hit* scores was above 50%. This, however, was not based on their d prime scores but rather on the percentage of their correct responses to the different pairs. This still indicates that FDHs' showed learning of some of the contrasts.

Group		Response: Different	Response: Same	Total
FDH	Pair: different	H= 467	M= 253	720
	Pair: same	F= 53	CR= 346	399
NS	Pair: different	H=355	M=5	360
	Pair: same	F=12	M=187	199

Table 8.1 Descriptive statistics of the responses the foreign domestic helper (DH) and NS groups provided in the AX discrimination task (CR=correct rejection, F=false alarm, H=hit, M=miss)

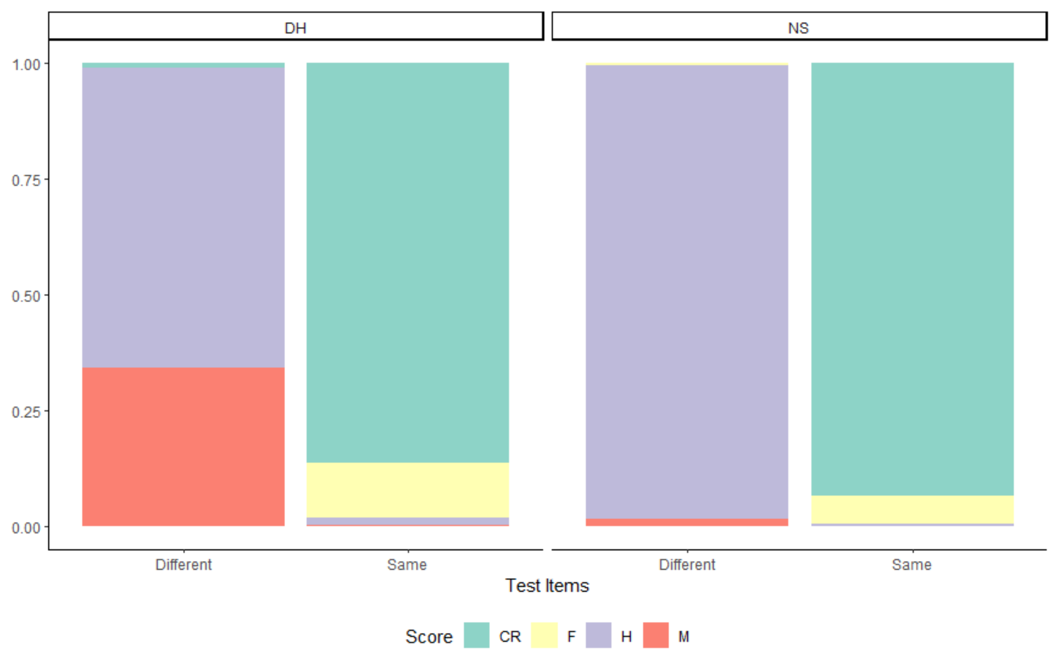


Figure 8.1 Percentages of responses FDHs and NSs provided in the AX discrimination task (CR=correct rejection, F=false alarm, H=hit, M=miss)

Table 8.2 shows further descriptive statistics that were obtained from a number of operations using results from table 8.1. The accuracy of the FDHs was 0.72, while that of NSs was 0.96. The error rate of FDHs (0.27) was higher than that of NSs (0.03). The true positive rate reflects the rate at which listeners responded ‘different’ when the stimulus was ‘different’. On the other hand, the false negative rate reflects the rate at which listeners responded ‘different’ when the stimulus was ‘same’.

Sensitivity (which reflects the rate at which listeners responded ‘same’ when stimulus was ‘same’) of both groups was higher that 50%. Precision reflects the rate at which listeners responded correctly when the stimulus was ‘different’. This was also higher than 70% for both groups. This confirms what I reported in Figure 8.1 that FDHs’ *hit* responses were above 50%

Group	Accuracy	Error	True positive rate (specificity)	False positive rate	True negative rate (sensitivity)	Precision
FDH	0.72	0.27	0.64	0.13	0.86	0.89
NS	0.96	0.03	0.98	0.06	0.93	0.96

Table 8.2 Descriptive statistics of the AX data

8.1.2 d Prime

A linear model revealed that the mean sensitivity score of the NS group (m=4.94) was significantly higher than that of the FDH group (m=2.17) (p<0.01; Table 8.3 and Figure 8.2). Taking a closer look at Figure 7.1, two of the FDHs (Si2 and Sw1) approached native-like discriminability of the consonant contrasts as they achieved sensitivity scores equal to those of some of the NS participants. However, their discrimination ability did not reach the mean score of the NS control group (See Appendix 9 for d-prime scores of all participants). Based on information from Chapter 5, these two FDHs had varying LoR in the Middle East (13 and 1.75 years, respectively), different Arabic literacy skills but comparable educational backgrounds. Thus, different factors could have contributed to their performance. Despite having comparable L1s with regard to the consonants of interest to the present study, Figure 7.1 shows that there is a large variability in the listeners’ d prime scores in this task as can be seen from the data points represented in the DH pink bar. As to what factors could have affected differences in performance, the coming section will discuss these in detail.

	Estimate	Std.Error	t value	Pr(>|t|)
DH (Baseline)	2.17	0.23	9.31	
NS	2.76	0.404	6.84	**1.93e-07 *****

Table 8.3 Linear model results of the effect of group (DH/NS) on the sensitivity scores obtained from the AX discrimination task

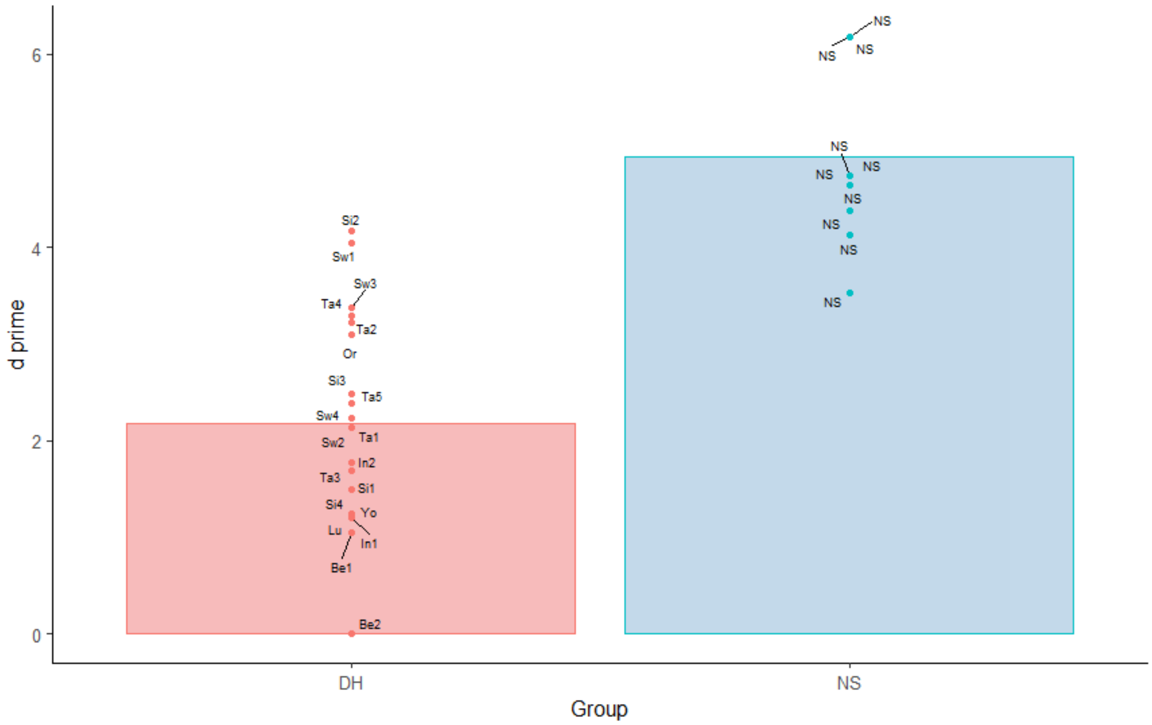


Figure 8.2 Mean *d prime* scores of the foreign domestic helper subjects (DH) and the native speaker (NS) control group

8.1.3 Factors affecting FDH's discriminability of consonantal contrasts

A linear model revealed that LoR and L2 literacy had no significant effect on FDH's discriminability of consonantal contrasts as Table 8.4 shows ($p>0.05$). Interestingly, FDHs' L1 education (years of formal schooling) as a continuous variable appeared to play a significant role in FDH's sensitivity scores ($p<0.05$). The effect of each of these factors will be discussed in detail in the following sections.

Predictor	Estimate	Std. Error	t value	Pr(>|z|)
(baseline)	0.55	0.65	0.85	0.404
LoR	0.02	0.03	0.75	0.46
L1 education	0.18	0.06	2.78	**0.01***
L2 Literacy	0.25	0.44	0.56	0.57

Table 8.4 Results of the linear model for the effect of LoR and L1 and L2 literacy on FDHs phonological sensitivity

7.1.1.1 LoR

Figure 8.3 further illustrates that there is no linear relationship between LoR and sensitivity scores of the FDH participants. Thus, no matter how long the FDH had spent in the Arabic-speaking world, her phonological sensitivity to consonant contrasts had not changed or improved. It is essential to note though that LoR only represents the period that a FDH had spent in the Arabic-speaking world and does not indicate actual input or interaction in the target language a given learner had had.

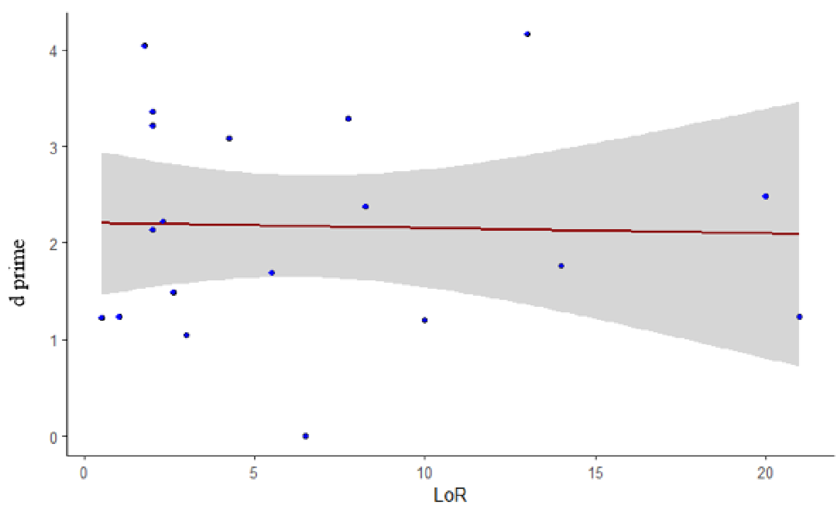


Figure 8.3 The effect of LoR on FDH's phonological sensitivity

7.1.1.2 Years of L1 Schooling

Figure 8.4 shows a positive, linear relationship between years of schooling and sensitivity scores of FDHs, confirming the significant effect of this variable on sensitivity scores as the linear model revealed. The more years the FDH had spent at school in her native language, the greater her discriminability of the contrastive sounds in Arabic was. There are three outliers that did not fit within this trend and can be considered as exceptions.

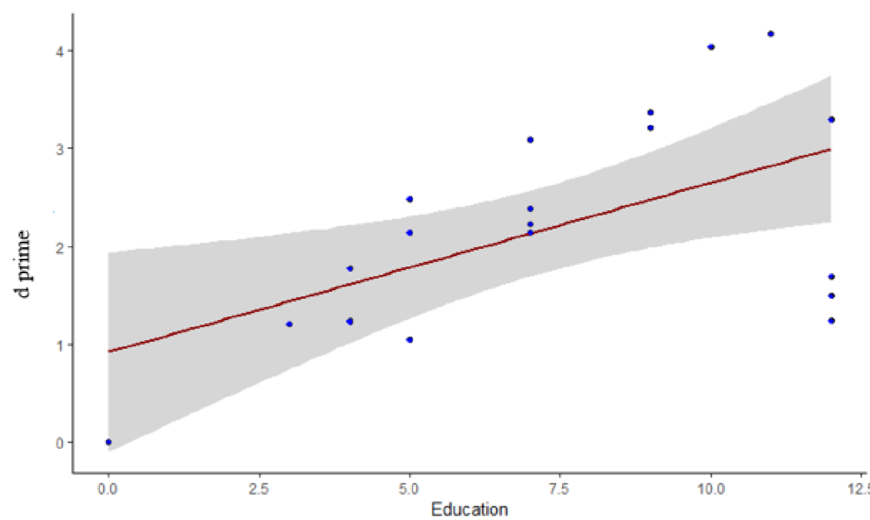


Figure 8.4 The effect of L1 formal schooling on FDHs phonological sensitivity

7.1.1.3 L2 literacy

Figure 8.5 illustrates that the mean sensitivity score of FDHs who were literate in Arabic is higher than that of FDHs who were non-literate in Arabic. This difference in performance based on Arabic literacy was not significant though. To recall from Chapter 5, non-literate FDHs include both Muslim and non-Muslim participants. Since literacy in Arabic can be associated with Islam in terms of worship using that language, data points of the Non-Muslim participants have been defined in the graph to show their sensitivity scores compared to literate and non-literate Muslim participants. These show that exposure to Arabic via Islam does not necessarily mean that the learners' performance would be better, as some non-Muslim subjects performed as well as the Muslim subjects.

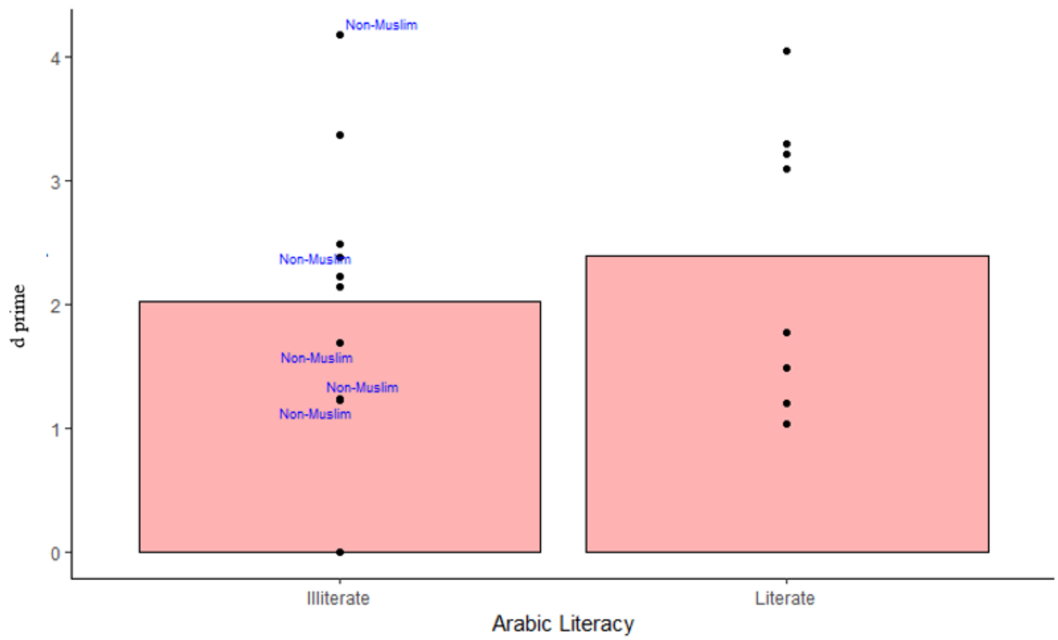


Figure 8.5 The effect of L2 literacy on FDHs phonological sensitivity

8.1.2 Summary and discussion

Results from this section provided support to H5 which predicted that FDHs' discriminability pf Arabic phonemic contrasts will not reach native-like performance. It was found that FDHs fell short of nativelike sensitivity to Arabic phonemic contrasts when their performance was compared to the NS control group. This provides support to research that proved the difficulty adults learning a second language encounter in perceiving L2 contrasts that are not phonemic in their L1s (Best and Strange 1992; Yamada 1995; Iverson and Kuhl 1996). Also, results provided partial support to H6. Native language formal education contributed significantly to FDHs' sensitivity scores. The longer a FDH had spent at school in her native language, the higher her sensitivity score was. Moreover, FDHs' who were literate in Arabic outperformed those who were non-literate in Arabic in the discrimination task. However, the contribution of L2 literacy was non-significant. LoR also did not play any significant role in FDHs phonological sensitivity. I highlighted earlier that two FDHs (Si2 & Sw1) had native-like performance based on their perceptual sensitivity scores. Both FDHs varied on several linguistic and social aspects clarified in Table 8.5. They had different L1s, AoA, LoR and L2 literacy skills but were comparable in terms of the period of time they had spent in formal schooling. Hence, L1 literacy appears to explain their comparable sensitivity to Arabic phonemic contrasts. Further discussion on the interpretation and implications of findings from this section will be presented in Chapter 8.

	Si2	Sw1
L1	Sinhala	Swahili
Age at Testing (AaT)	46	19
Age of Arrival (AoA)	33	16
Length of residence (LoR)	13	1.75
L1 formal schooling	10 years of formal schooling	10 years of formal schooling
L2 literacy	none	Basic reading and writing skills

Table 8.5 Background information of the two participants who performed like NSs in the AX task

8.2 FDHs' Production of Arabic Consonants

This section provides results on FDH's accuracy at producing marked Arabic consonants. This examination was set to test the following hypotheses:

> **H7.** FDH will fall short of target-like accuracy in producing Arabic consonants.
>
> **H8.** FDHs with longer LoR, and who have more formal schooling and who are literate in their L2 will be more accurate at producing consonants.

Figure 8.6 demonstrates that the dentals, the emphatics, the pharyngeals and the uvulars were not produced as accurately as sounds such as /k/, /s/ and /t/. The percentage of correct productions of former consonants was below 50, while that of /k/, /s/ and /t/ was above 90.

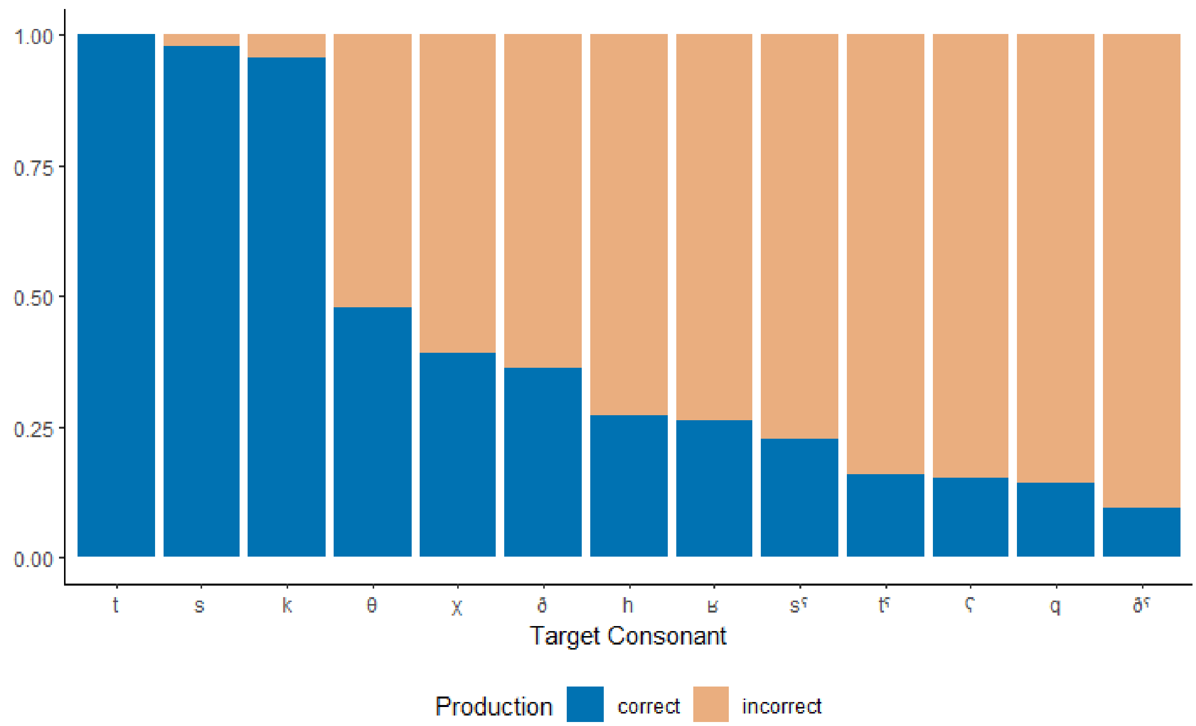


Figure 8.6 Percentage of overall correct and incorrect productions averaged across target consonants

Figure 8.7 illustrates the replacements used by FDHs when their production of the target consonants was considered incorrect or inaccurate. It also shows the proportion of the correct production, as well as the proportion of the replacement used in cases of incorrect productions, averaged across target consonants. All speakers who produced /ð/ incorrectly realized it as [d]. The Oromo speaker also realized /ð/ as [z]. When /ðˤ/ was produced incorrectly, it was more frequently realized as /ð/ but also as other consonants including [d], [m], [dˤ] and [s]. The voiced uvular /ʁ/ was most often realized as [g] but also as [k], [h] and [q]. In some cases, /ʁ/ was deleted.[23] Although /k/ was most often produced correctly by the FDHs, in a few cases it was realized as [ʃ] or [tʃ]. Substitutions of /k/ are likely to be variants of this consonant used in other Arabic dialects; for example, /k/ is realized as [tʃ] in some Gulf Arabic dialects as indicated in Chapter 4. As explained in Chapter 5, some FDHs had lived and worked in some Gulf Arab countries before moving to Oman. Speakers who produced /k/ as a Gulf Arabic variant were In1, an Indonesian speaker who had worked in Saudi, UAE and Kuwait before migration to Oman, and Si3, a Sinhala speaker who had worked in Saudi and Kuwait before migration to Oman. Speakers who produced /χ/ incorrectly most often realized it as [q] and less often as [k], [h] and [g]. In a few cases, /χ/ was deleted. In most cases, /q/ was realized as [k], but in a few cases it was realized as [g], [t] and [χ]. When /ʕ/ was produced incorrectly, it was most often deleted or realized as a glottal stop /ʔ/. In a few cases, it was realized as its voiceless counterpart [h]. Although /s/ was most often produced correctly by the FDHs, in one case it was produced as [ʃ] by a Bengali speaker. The emphatic fricative /sˤ/ was most frequently deemphasized and realized as [s]. In a few cases, it was also realized as [ʃ], and this was by the same Bengali speaker who realized /s/ as [ʃ]. When /tˤ/ was produced incorrectly, it was deemphasized and realized as [t] in all productions attested. /t/ was the only sound produced correctly in 100% of the cases. Finally, /θ/ was most often realized as [t] and less often as [s] (by a Sinhala speaker) or [tˤ] (by a Bengali speaker).

[23] Deletion of a sound is represented by the symbol ∅ in the chart which also has the red color coding.

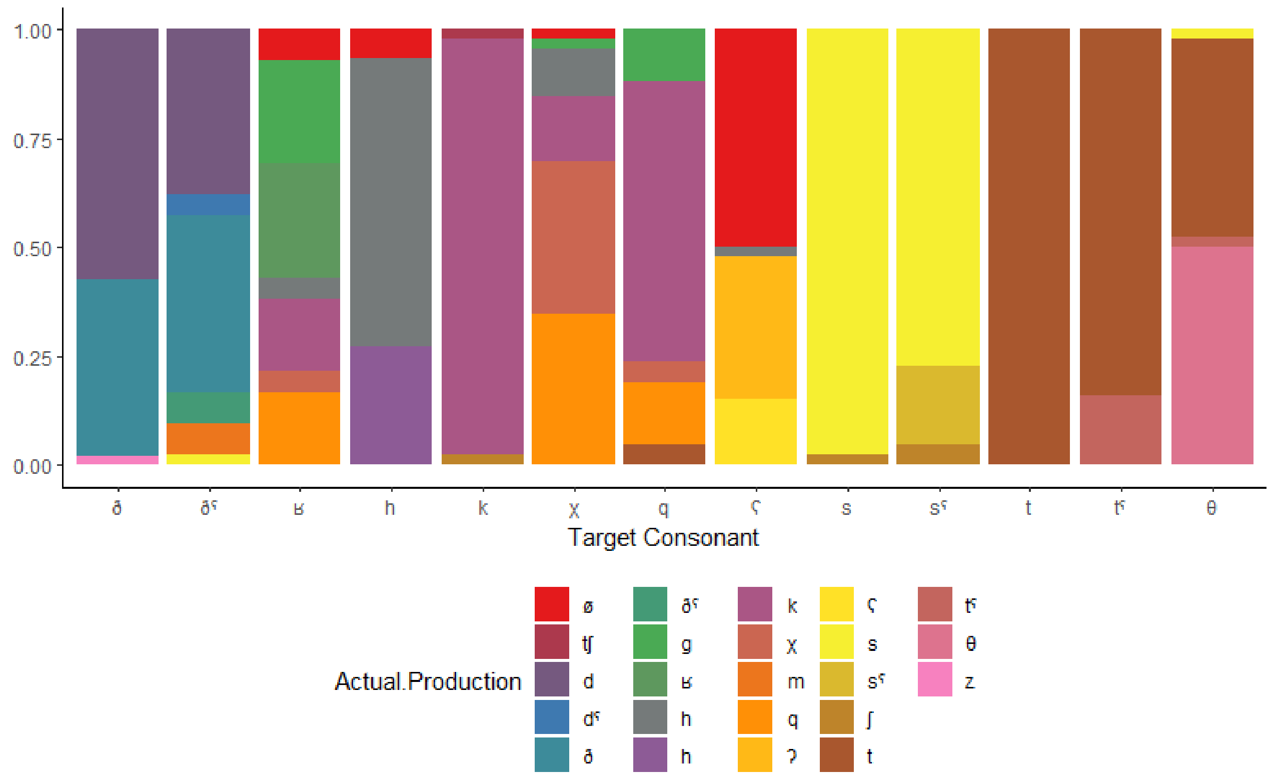


Figure 8.7 Proportions of correct and incorrect productions of target consonants represented by the realizations used by the speakers.

8.2.1 Factors affecting the production of Arabic consonants

A GLMM demonstrated the absence of significant effects of LoR and native language schooling on the accurate production of Arabic consonants by FDHs ($p>0.05$) (Table 8.6). L2 literacy proved to have a significant effect on FDHs' productions of the target sounds ($p<0.05$).

Predictor	Estimate	Std. Error	Z value	Pr(>|z|)
(baseline)	0.14	0.28	0.49	0.62
LoR	0.02	0.01	1.41	0.15
Native Language Schooling	0.03	0.03	0.98	0.32
L2 Literacy	-0.53	0.22	-2.34	**0.019 ***
Random effects Variance: 0.09		Std. Deviation: 0.31		

Table 8.6 Results of the GLMM for the effect of LoR and L1 and L2 literacy on FDHs production accuracy

8.2.1.1 LoR

Similar to findings on the discriminability of consonant contrasts, LoR did not play any role in how accurate the FDHs were at producing the target consonants as illustrated in Figure 8.8. In fact, speakers with the longest LoR appeared to be less accurate (more foreign-accented) than those with shorter length of residence, as the trend in Figure 8.8 shows a slightly sharper drop towards LoR of 18 years and more.

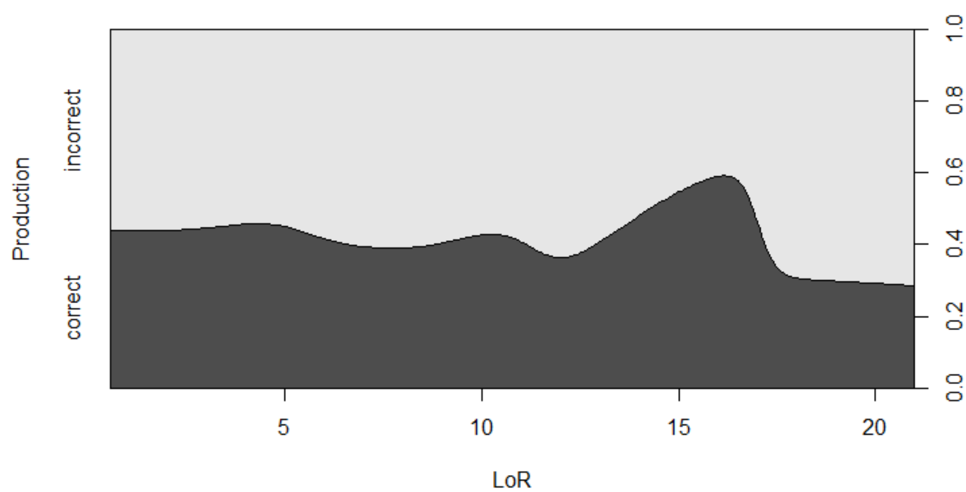


Figure 8.8 The effect of LoR on FDHs' production of consonants

8.2.1.2 L1 schooling

Despite its significant effect on FDHs' sensitivity of consonantal contrasts, native language schooling did not have any impact on participants' production accuracy as shown in Figure 8.9. The trend is almost equally constant along years of formal education.

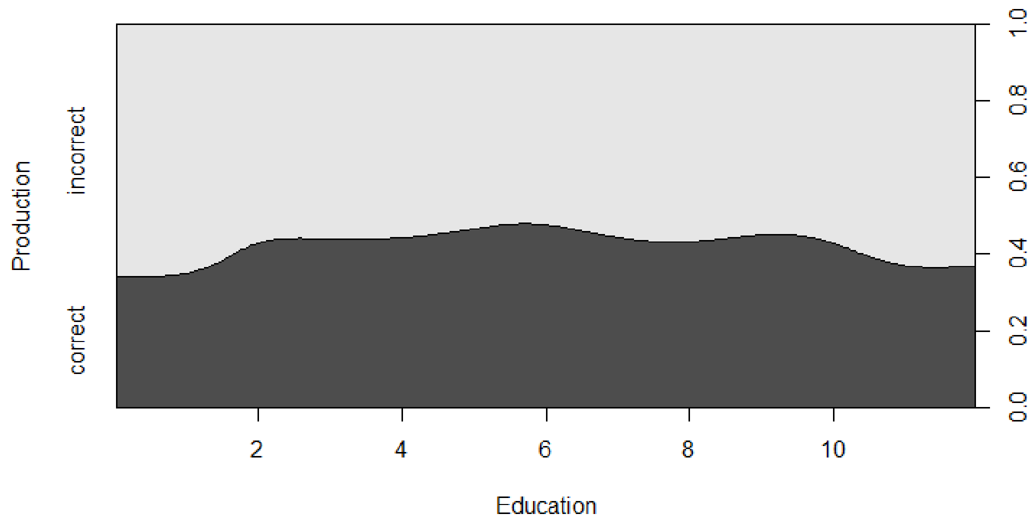


Figure 8.9 The effect of native language literacy (education) on FDHs production of consonants

8.2.1.3 L2 literacy

Figure 8.10 shows that speakers who were literate in Arabic had more correct productions (50%) of the target consonants than those who were non-literate (37.8%). This difference was significant according to the GLMM in Table 8.6. This means that FDHs who learned Classical Arabic via listening to Arabic and reading it during recitation of the Qur'an were less foreign accented compared to those who did not learn to read Arabic (whether they were Muslims or non-Muslims).

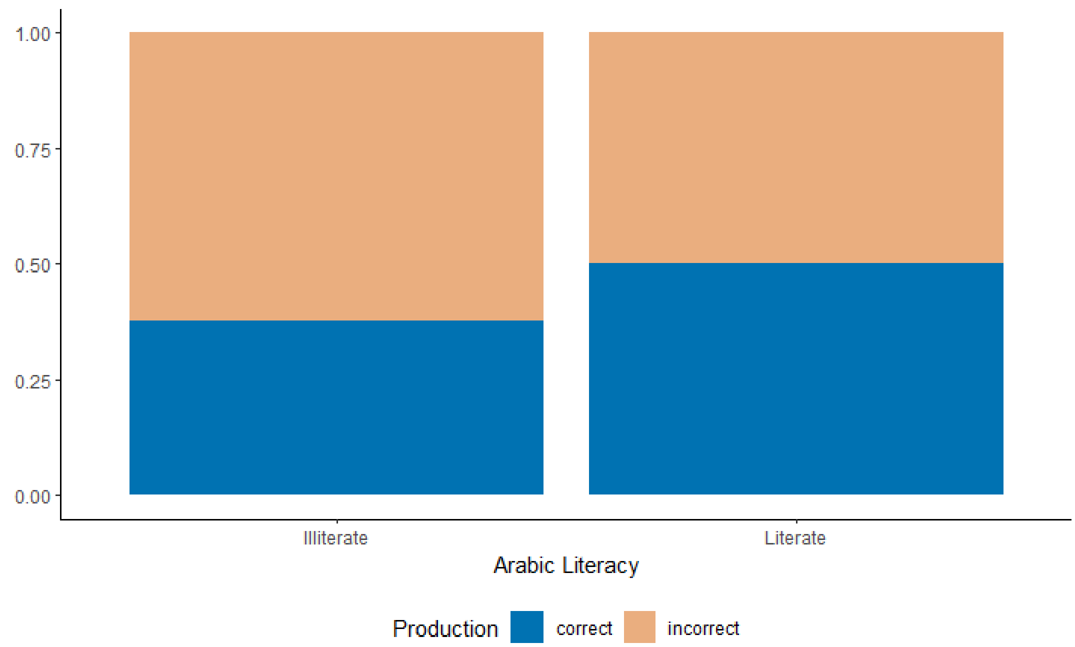


Figure 8.10 The effect of L2 literacy on FDHs production of consonants

8.2.2 Summary and discussion

This section provided results concerning the extent to which FDHs were accurate at producing complex Arabic consonants and the factors that were most likely to affect their differential performance. Results from this section confirmed H7 and demonstrated that FDHs' did not reach target-like proficiency in their production of complex consonants. The sound substitutions they used when producing these consonants were suggestive of a perceived foreign accent. H8 was also partially supported by results on the factors that influence FDHs performance in producing consonants. Arabic literacy showed a robust effect on FDHs' accuracy of consonant productions, supporting my previous hypothesis. This could be attributed to the age they started learning Arabic. However, LoR and L1 schooling did not play a significant role in FDHs' production accuracy contrary to my expectations. Further discussion on these will be provided in Chapter 9.

8.3 Production of Consonant Clusters

Results from this section will either confirm or reject the following hypotheses:

> **H9.** FDH will fall short of target-like accuracy in producing Arabic consonant clusters.
>
> **H9.**a FDHs will be more successful at producing onset consonant clusters than coda consonant clusters due to the markedness of the latter.
>
> **H10.** FDHs with longer LoR, and who have more formal schooling and who are literate in their L2 will be more successful at producing Arabic consonant clusters.

Table 8.7 shows that there is a greater tendency by FDHs to retain the short vowel in initial unstressed syllable structures (89.4%) than produce onset consonant clusters.[24] To clarify further, taking into consideration the optionality of consonant clusters in the onset of the target dialect, FDHs tended to use the less marked structure, one that would result in a CV syllable, more frequently. This indicates that FDHs have not acquired complex onsets despite their exposure to them in the target input as was evident from results on FDH-directed speech in Chapter 6. Relatively equal chances of modification and non-modification of codas were present in FDHs' performance (Table 8.7). Percentage of modification of coda clusters, generally, show that FDHs have not yet reached target-like performance of these clusters.

Syllable Type	**Syllable Modification**	
	Yes	**No**
Onset Consonant Clusters	(118) 89.4%	(14) 10.6%
Coda Consonant Clusters	(43) 48.9%	(45) 51.13%

Table 8.7 Percentage of modification in onset and coda consonant clusters

[24] 'Syllable Modification', as stated in the table, might not be the right term to describe the subjects' production of onsets with optional clusters, provided that native speakers show variation in realizing these consonant clusters (either by retaining the vowel or deleting it). However, for ease of description especially in graphs and tables, the term 'modification' will be used to describe the subjects' production pattern of this aspect of the syllable. Hence, 'modification' in this sense implies that CCs in the onset are the standard in the target dialect.

The tendency to modify or break coda consonant clusters (48.9%) is less frequent than the tendency to modify onset consonant clusters (89.4%). A GLMM demonstrated that this difference was significant (See Table 8.8 and Figure 8.11).

Predictor	**Estimate**	**Std. Error**	**Z value**	**Pr(>|z|)**
WF cluster (baseline)	0.046	0.31	0.14	0.88
WI cluster	-2.51	0.41	-6.07	**1.25e-09 *****
Random effects:	Variance: 0.94	Std. Deviation: 0.96		

Table 8.8 Results of GLMMs showing the difference between consonant clusters word initially and word finally in relation to modification

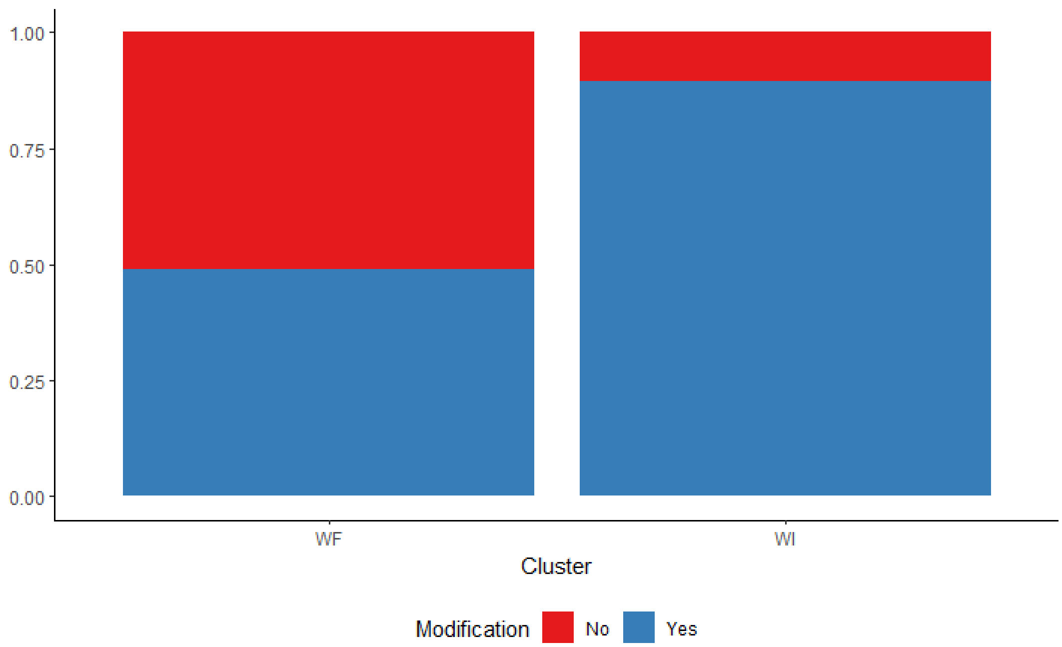


Figure 8.11 Percentage of onset and coda consonant clusters modification

Table 8.9 and Figure 8.9 illustrate that the main strategy the speakers used to modify the complex syllable structure for both onset and coda clusters was vowel epenthesis (98.3% and 100%, respectively). There were only two cases of consonant deletion and that was in onset clusters.

Syllable Type	**Modification Strategy**	
	Consonant Deletion	**Vowel Epenthesis**
Onset Consonant Cluster	(2) 1.69%	(116) 98.305%
Coda Consonant Cluster	0	(43) 100%

Table 8.9 Percentages of the strategy used in modifying initial and final CCc

8.3.1 Factors affecting the production of onset consonant clusters

8.3.1.1 L1

The trend in Figure 8.12 demonstrate that, generally, the tendency to produce less marked onset consonant clusters is evident in all FDHs' productions regardless of the L1. As Figure 8.12 shows, above 50% of FDHs' productions involve modifications of the onset. FDHs with Oromo, Sinhala and Telugu L1 backgrounds had the highest rate of onset modification, which also means that they had no instances of onset clusters in their productions and always produced a less marked, CV syllable. FDHs with Indonesian, Tagalog, Bengali, Yoruba, Swahili and Luganda L1 backgrounds showed variation in their modification of onsets. Throughout their productions of the target words, they sometimes deleted the short vowel and maintained the consonant clusters and other times retained the vowel and produced a CV syllable just like NSs. To illustrate, these FDHs produced the target word /ktaːb/ sometimes as [ktaːb] and others as [kɪtaːb]. The Luganda native speaker had the highest rate of onset clusters (33.33%).

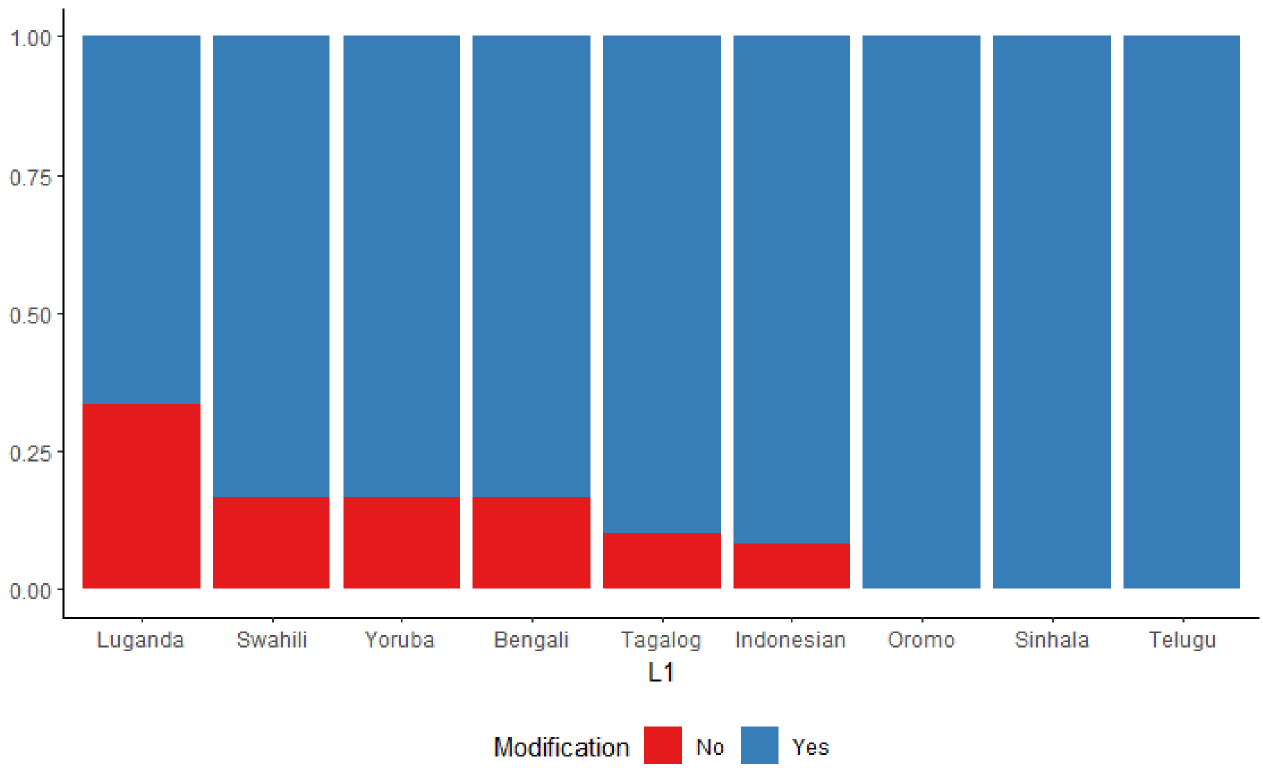


Figure 8.12 The effect of FDHs' L1 on CC modification in onsets

8.3.1.2 LoR

Figure 8.13 shows the effect of LoR on FDHs' productions of onset clusters. As we can see, the general trend shows that the deletion of the vowel (Yes) is far more frequent than the retention of the vowel (No). Less modification of the onset appears in productions of FDHs who had the shortest LoR. The trend then shows a stable change between those of 5 years and 15 years of LoR before it starts dropping. The drop indicates that those with the longest LoR had the most frequent modifications to the syllable. From this, we can conclude that LoR does not play any role on whether onset consonant clusters will be maintained or simplified in FDHs' productions.

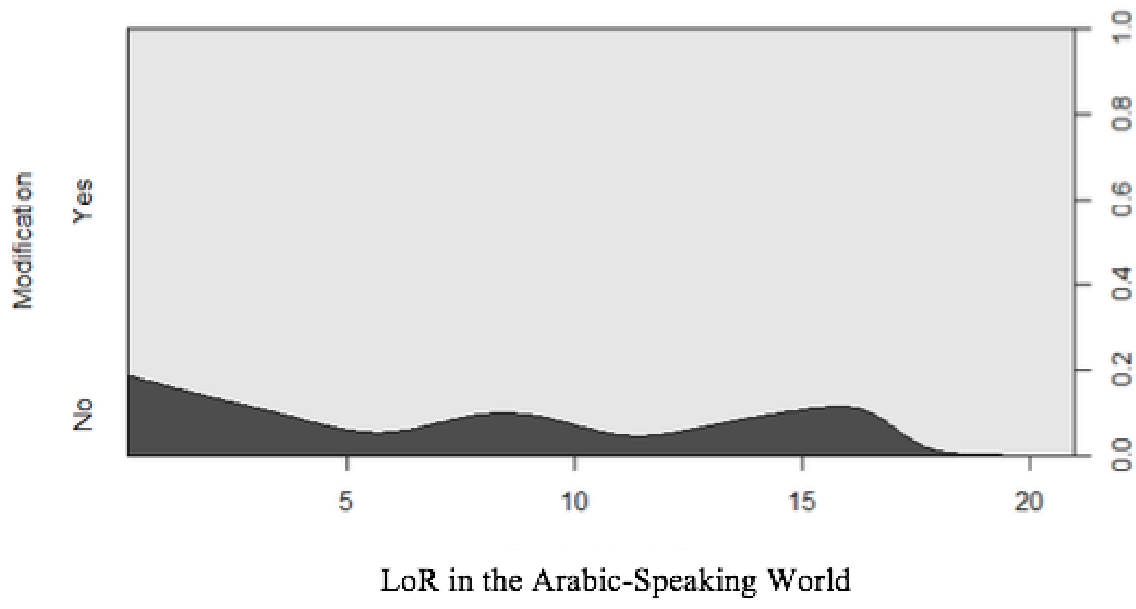


Figure 8.13 The effect of LoR on modification of onset CCs

8.3.1.3 L1 schooling

Modification of initial consonant clusters do not appear to change considerably in relation to years of formal schooling. Figure 8.14 shows a stable pattern of onset cluster productions regardless of years of schooling. Slightly more production of consonant clusters appeare in FDHs with LoR shorter than 2 years.

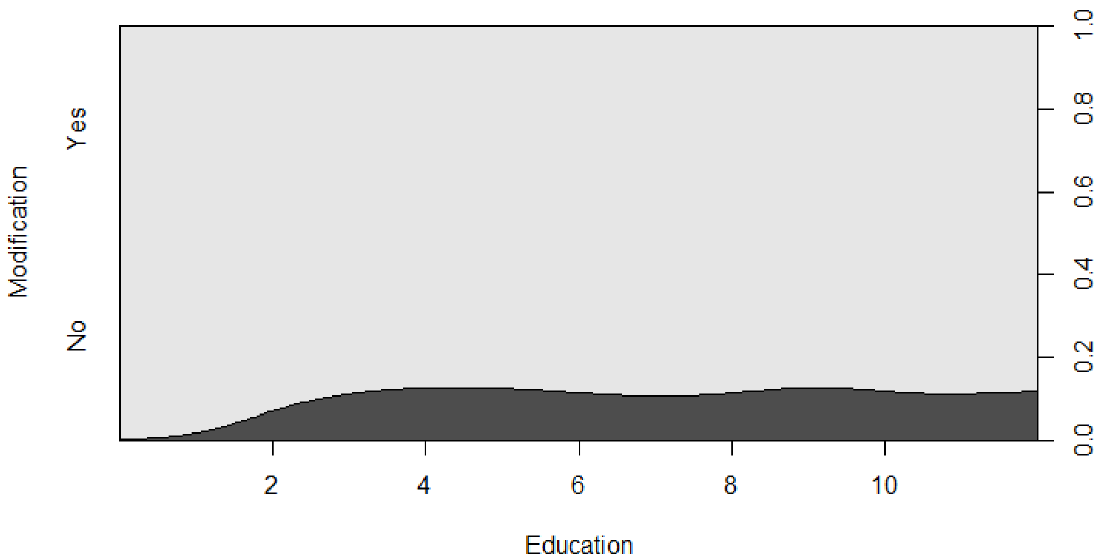


Figure 8.14 The of effect of L1 schooling on FDHs' modification of onset CCs

8.3.1.4 L2 literacy

Figure 8.15 illustrates that FDHs who were literate in Arabic exhibited less modification of onsets than FDHs who were non-literate in Arabic. Differently stated, literate FDHs produced more onset consonant clusters (16.66%) than non-literate FDHs (7.77%).

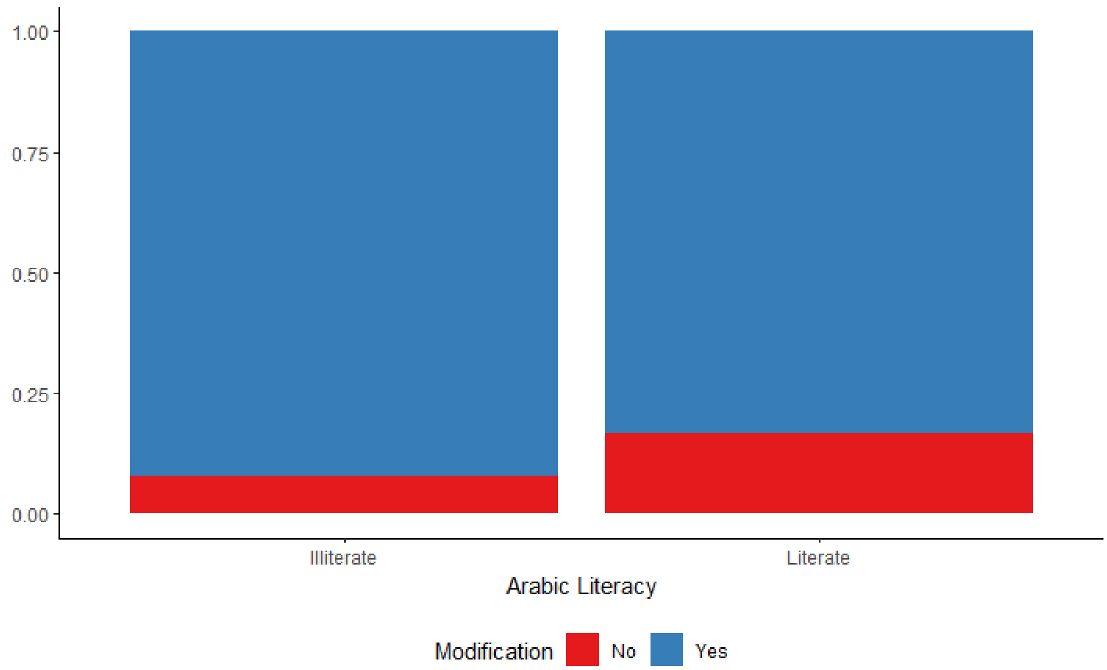


Figure 8.15 The effect of Arabic literacy on modification of onset CCs

8.3.2 Factors affecting the production of coda consonant clusters

A GLMM revealed that LoR had a significantly negative effect on the modification of the consonant clusters (Table 8.10). The more FDHs had spent in the Arab world the less they were successful at producing coda consonant clusters. L1 years of formal education appeared to have a significant effect on the modification of codas. However, L2 literacy did not play any significant role in the pattern of consonant cluster production. Each of these factors will be discussed in detail in the following sections.

Predictor	Estimate	Std. Error	Z value	Pr(>|z|)
(baseline)	-0.56	0.64	-0.88	
LoR	-0.102	0.04	-2.12	**0.03 ***
Native Language Schooling	0.13	0.06	1.96	**0.04 ***
L2 Literacy	0.85	0.55	1.54	0.12
Random effects	Variance: 0.08		Std. Deviation: 0.28	

Table 8.10 GLMM results on the effect of cognitive and social factors on production on final CCs

8.3.2.1 LoR

Figure 8.16 shows the effect of LoR in the Arab world on FDHs' productions of coda consonant clusters. Success at producing coda consonant clusters correlated negatively with LoR. The pattern shown in Figure 8.16 suggests that as LoR extends, FDHs produce less native-like syllables. FDHs with the shortest LoR produced coda clusters (native-like performance) more frequently than those with longer LoR. The trend also shows fluctuation which implies existing individual differences in how frequently coda clusters were produced with regard to LoR.

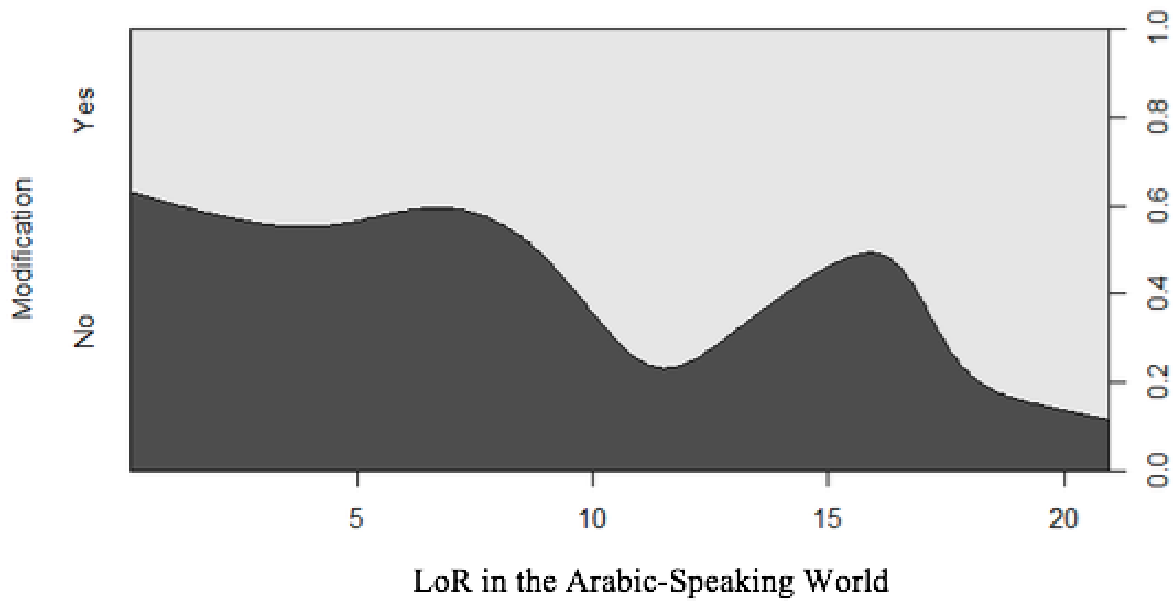


Figure 8.16 The effect of LoR in Arabic-speaking countries on modification of codas

8.3.2.2 L1 schooling

Interestingly, Figure 8.17 shows a positive, linear relationship between years of formal education and success at producing coda clusters. The more educated a FDH was, the more successful she was at producing a native-like syllable structure. This is in line with the significant role L1 schooling played in FDHs' performance in the discrimination task.

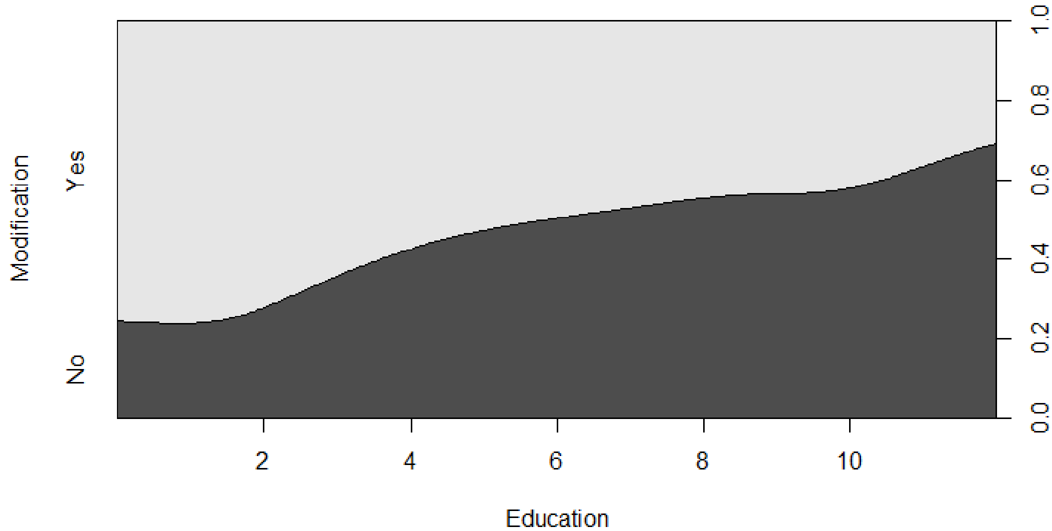


Figure 8.17 The effect of L1 schooling on modification of coda CCs

8.3.2.3 L2 literacy

Although L2 literacy played no significant role in FDHs' production patterns based on GLMM, Figure 8.18 illustrates that FDHs who reported they were literate in Arabic performed better at producing coda consonant clusters than those who reported being non-literate. Literate FDHs produced coda clusters in 64.28% of the instances while non-literate FDHs produced them in 45% of the instances. Hence, literacy in Arabic, or more specifically knowledge of the Arabic script seems to aid L2 learners' acquisition of native-like oral forms.

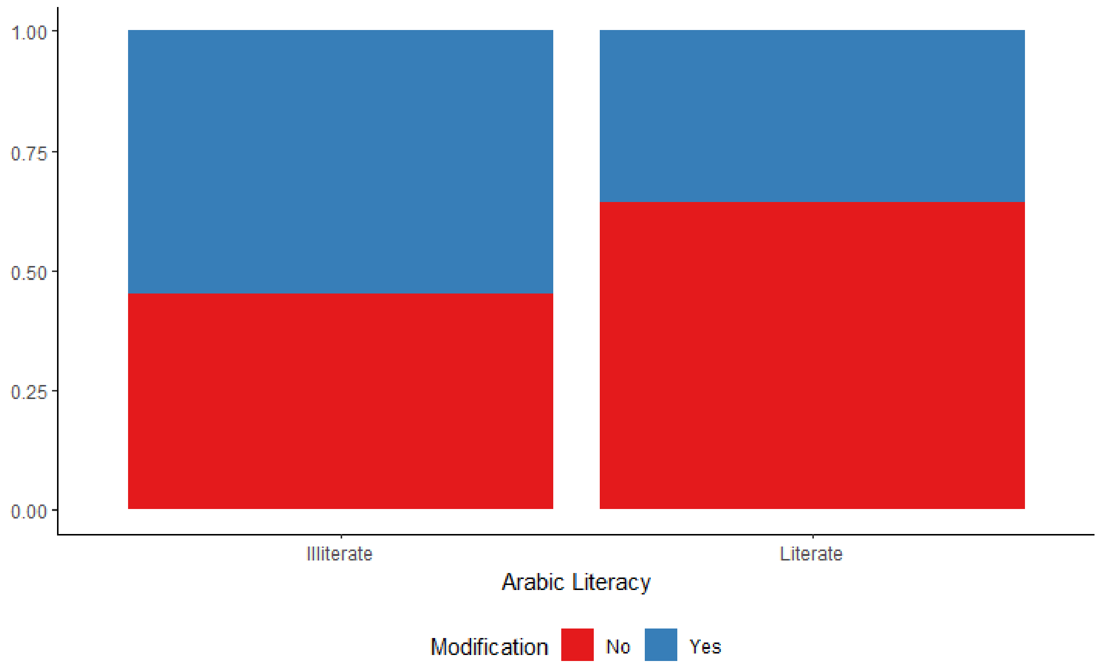


Figure 8.18 The effect of Arabic literacy on modification of final CCs

8.3.3 Summary and discussion

Results on FDHs production of consonant clusters confirmed H9 and demonstrated that FDH fell short of target-like accuracy in producing Arabic consonant clusters. Results on FDHs production of onset and coda clusters did not provide suport to H9a and showed that FDHs' productions contained significantly more coda clusters than onset clusters. This finding does not provide support to the universality of language principles and markedness principles. In Chapter 2, we learned that coda clusters are more marked than onset clusters. FDHs' production patterns, however, indicate that they were more successful at producing complex codas of Arabic clusters than complex onsets. This can be explained in terms of the NS input these learners are exposed to (See Chapter 9).

As to the factors that play a role in the learners' differential performance, H10 was partially supported for both onset and coda clusters. The L1s of FDHs appeared to play some role in their production patterns of onset clusters, supporting the prediction that FDHs' L1 will influence their production of onset clusters. FDHs whose L1s allow onset clusters produced more of these than those whose L1s lack them (except for the Yoruba and Indonesian speakers whose L1s do not allow onset clusters but still showed some productions of Arabic onset clusters). LoR and L1 schooling did not show any significant effects on FDHs' production pattern of onset clusters. L2 literacy, however, appeared to play some role. FDHs with knowledge of Arabic script produced more onset clusters compared to those who did not have that knowledge. For coda clusters, LoR and L1 formal schooling played a significant role in FDHs' production patterns. The longer FDHs had lived in the Arab world, the less they produced coda clusters. This was against my prediction that as FDHs' LoR increases, their production will contain more consonant clusters (native-like productions). The effect of L1 schooling was in line with my hypothesis: the longer FDHs' had spent at school the more consonant clusters their productions would contain. Similar to onset clusters, coda clusters were produced more often by FDHs who were literate in Arabic though this trend was not significant.

Chapter 9. Discussion

9.0 Introduction

This Chapter will first discuss the linguistic adaptations found in FDH- directed speech with regard to their implications to the wider research. It will specifically highlight the phonological, acoustic and prosodic changes attested in FDH-directed speech and compares them to those reported in IDS, clear speech and FDS elaborating on the role of the interaction context on variation in the output. Whether or why hyperarticulation is evident in FDH-directed speech will be justified after that. This will be followed by interpretations for the role of FDHs' foreign accentedness, LoR and religion in adaptations attested in FDH-directed speech. Then, the Chapter will discuss the implication and interpretations of the second part of the study, which is related to the perception and production of Arabic consonants and consonant clusters. It will highlight the role of cognitive (L1, L1 schooling, L2 literacy) and social factors (LoR) in the participants' performance and will elaborate on this in relation to theories of L2 speech learning.

8.1 Interpretation and Implications of the Linguistic Adaptations in FDH-Directed Speech

To investigate the linguistic adaptations in FDH-directed speech, the following research questions were put forward for examination:

> **Q1)** What (if any) systematic adaptations are made in consonant, consonant cluster and vowel productions in FDH-directed speech? To what extent are these adaptations similar to or different from those reported in IDS, FDS, and clear speech research?
>
> > **Q1a)** Does FDH-directed speech contain hyperarticulation/enhancement of sound contrasts?
> >
> > **Q1b)** Does LoR play a role in any adaptations made in speech to FDHs?

Answers to these questions as obtained from the results of the study will be provided in the following sections along with a discussion of the findings.

8.1.1 Phonological adaptations in FDH-directed speech

Phonological examination was set to test the following hypotheses:

H1. NSs will simplify complex consonants and produce sounds that are familiar to FDHs (e.g. /ʁ/→/g/; /q/→/k/).

H2. FDH-directed speech will involve less marked onset and coda consonant clusters.

Interestingly, the auditory analysis of complex consonants yielded no differences between FDH-directed speech and ADS with regard to the realization of complex consonants. Acoustic profiling of all the consonants examined showed no difference in the acoustic patterning of consonants between FDH-directed speech and ADS. Hence, H1 was not supported as NSs did not accommodate to their FDHs interlanguage. Findings from consonant cluster data pointed to a partial support to H2. FDH-directed speech was characterized with more simplification of onset consonant clusters compared to ADS. For example, the target word /**dg**aːg/ 'chicken' was realized more often as [dagaːg] in FDH-directed speech compared to ADS. Contrary to this finding, coda consonant clusters did not undergo any simplification in NSs interactions with FDHs. The following paragraphs will provide a discussion of these findings.

In an interaction setting where a native speaker (NS) is involved in a conversation with a non-native speaker (NNS) such as the one in the present study, adjustments of speech are expected on the part of the NS (Ferguson 1975; Ferguson 1981; Long 1983; Uther *et al.* 2007; Rodriguez-Cuadrado 2018). When the second language (L2) learner interlocutors are a population of immigrants perceived by NSs as inferior, linguistically limited and speaking with a foreign accent such as the FDHs in the current study, it is expected that the NSs would use ungrammatical speech (including incorrect word order), simpler constructions, pidginized speech and/or repetitions of their errors (Ferguson 1975; Clyne 1981; Avram 2014; Alfallaj 2016). One possible scenario in such a linguistic situation is that NSs would adjust their marked consonants and consonant clusters for the purpose of simplification or accommodation to the interlocutors' interlanguage. Results on FDHs' perception and production of marked Arabic consonants provide a better picture of these learners' interlanguage system. We learned from Chapter 7 that FDHs' sensitivity scores were significantly lower than those of the NS control group. This indicates that FDHs perceive Arabic consonant contrasts in a different way than how NSs perceive them. Moreover, FDHs' production accuracy rate of complex Arabic consonants was less than 50% as shown in Chapter 7. Evidence for this was shown in the substitutions FDHs used to produce the target consonants. Common substitutions used are shown in Table 8.1.

Target	/θ/	/χ/	/ð/	/ħ/	/ʁ/	/sˤ/	/ʕ/	/q/	/ðˤ/
Substitution	[t]	[q]	[d]	[h]	[g]	[s]	deleted	[k]	[ð]

Table 9.1 Common substitutions used among FDHs in the production of Arabic consonants

From this, we have evidence that (1) FDHs' perceived Arabic consonant contrasts in a way different than that of NSs, though this is not accessible to the NSs, and (2) FDHs had more frequent inaccurate productions of the examined consonants, which is accessible to NSs. One might therefore expect NSs to modify their speech to the FDHs accordingly, in response to their interlanguage production. However, results from auditory analysis and acoustic profiling in the current study revealed that complex Arabic consonants undergo no adjustment in the speech of the NS participants who were involved in an interaction with their FDHs. In other words, NSs in this study produced all the examined consonants in their target form and did not make any attempts to accommodate to their interlocutors by, for example, producing consonants that are available in their interlocutors' interlanguage system.

The aforementioned finding is in line with Smarts' (1990) observation that the phonology of the local Arabs who interact with foreigners is "preserved"; i.e. unchanged. Smart's account was not based on any embirical research and so we do not have further details that would be of benefit to the current study. Other than this, Smart reports that the structure of Gulf Pidgin, as he referred to it, is rather simplified. The few empirical studies that examined foreigner talk in a similar context as to the one in the present study demonstrate that NSs' speech to FDHs is morphologically and syntactically modified for the purpose of providing a more simplified linguistic structure that is more familiar to their interlocutors (Bizri 2009; Dashti 2013). The current study focused instead on phonology and examined segmental and syllable structure modifications. Additionally, the lack of accommodation to the FDHs' interlanguage is not in line with what other studies reported regarding speakers' accommodation to their addressees' language errors (e.g. Clyne 1981; Katz 1981; Evans 1987). For example, Clyne (1981) reports that second generation immigrants accommodated to their non-native English parents' segmental errors when producing Australian English. English sounds that did not exist in the parents' L1s were replaced by sounds from their native languages (e.g /d/→[t], /ð/→[d] or [s], /θ/→[t]). Clyne explains this accommodation as a result of speakers' attempts to identify with their addressees. Findings from the current study diverge from those of Clyne. This can be explained in light of the Communication Accommodation Theory (CAT) (Chapter 3). The non-use of foreigner talk with regard to segmental adaptations in the current study can be interpreted by relating it to how the NS employers perceive their FDHs. Unlike the situation in second generation foreigner talk, the NSs in the current study did not feel a need to identify with their FDHs due to the existence of such role differences as high-status-low-status and non-foreign-foreign as well as power relations such as employer-domestic helper. The absence of segmental adaptations reflects the strategies of *complementarity* or *maintenance* in which the NSs may or may not adjust their speech based on how they perceive role differences with their interactants. NSs might have wished to emphasize the non-foreign-foreign and employer-domestic helper role differences by maintaining their standard dialect (Janicki 1986; Zuengler 1991).

A common belief in the literature about foreigner talk is that it serves to maintain the social distance between NSs and NNSs who are considered subservient or inferior especially when simplification and reduction are used in speech addressed to them (Valdman 1981). Using simplifying processes implies that NSs "equate social inferiority with intellectual limitation" (Valdman 1981:44). Therefore, using foreigner talk denies NNSs input that would help them acquire the target language and from achieving acculturation which is a crucial psychological and social factor in second language learning (Schumann 1978; Valdman 1981). Simplifying speech sounds might reflect downgrading and mockery and if used might reflect the belief above (Valdman 1981). Although the perceived low status of the foreign addressee is considered a potential trigger for the use of extreme characteristics of foreigner talk (Zuengler 1991; Snow *et al*. 1981), NSs in the current study did not converge on their FDHs' errors when producing marked consonants. Hence, the absence of simplification of speech sounds in this study does not provide support for the belief that foreigner talk serves to maintain the social distance between NSs and NNSs through the use of simplifying processes and confirms that variability found in use of foreigner talk and its characteristics depends on the L2 learners' and interlocutors' context (Valdman 1981; Zuengler 1991). Valdman (1981) proposes that foreigner talk also features clarifying processes that may serve as a facilitative device and help enhance comprehension by providing the fullest forms of the target language and the use of redundant constructions. Additionally, the Interaction Hypothesis and studies conducted within its framework have mainly focused on how speech to NNSs is modified to aid comprehension and optimize communication through negotiation of meaning (Long 1981; Pica and Doughty 1985; Gass and Mackey 2007). Therefore, major findings of these studies described NS-NNS interactions to involve clarification requests, confirmation checks, self-repetitions, elaborations etc. (see Ellis 1991 for a review).

Despite the perceived lower status, perceived inferiority and limited Arabic competence of the FDHs not resulting in use of foreigner talk with regards to phonology, the NS-FDH context in the current study illuminats two-way interaction that is perhaps more aimed at achieving meaningful communication rather than focusing on adjusting or simplifying speech sounds. Communication between an employer and her FDH most often revolves around giving instructions and then making sure that the FDH understands her daily home chores. Hence, an employer will strive to make her message understandable by the FDH as will be shown later. NS-FDH interaction in the task they carried out for the current study involved similar communicative adjustments that the Interaction Hypothesis claim to be part of NS-NNS interactions. The following conversation between a NS and a FDH shows examples of such adjustments (the highlighted words are target words used in the task):

(8.1)

NS: fiː **fi**ː**l**? ħajwaːn ħajawaːn **fi**ː**l**. repetition and elaboration
(Is) there elephant? Animal animal elephant
“Is there an elephant? An animal, an animal, an elephant.”

FDH: (Silence) 2.33 seconds

NS: maː fiː? laː fiː. confirmation check
No there (is)? No there (is)
“There is none? Yes, there is.”

FDH: (Silence) 2.15 seconds

NS: laʔ ħajwaːn ħajwaːn. **fi**ː**l** **fi**ː**l**. repetition
No animal animal elephant elephant
“No, an animal, an animal. An elephant, an elephant.”

FDH: ajwa (Then says another word that was not audible)
Yes

NS: laʔ mɪθl mba: (goat sound) elaboration
No like baa
"No, it's like baa"

Such adjustments evident in FDH-directed speech show the NSs' effort to clarify her point by using repetition, confirmation check, and elaboration. This supports claims of the Interaction Hypothesis. It also shows the NSs struggle to make her point understandable to the FDH. These communicative adjustments, however, do not mean that linguistic simplification does not occur at higher prosodic levels (e.g. the syllable, the sentence, etc.). For example, I observed in data from the present study that NSs used simplifying processes at morphological or syntactic levels, but because the focus of the current study was on consonants and syllable structure per se, these were not subject to quantitative examination. Below, I list a number of simplifying processes used by NSs a long with examples from excerpts taken randomly from data collected for the current study:

(8.2) Article Deletion

a. muː loːn haːdhə Ø **korsɪ**?
What colour this chair?
'What colour is the chair?'

b. Fiː foːq Ø **tˤaːwleh** ħaliːb.
There on table milk.
'There is milk on the table?'

(8.3) Use of masculine gender marker only

a. Fiː sˤufrɪjjah **kabiːr**
(Is)there pot big.
'There is a big pot'

In (8.2), the Nizwa Arabic definite article /ʔɪl-/, which is frequently contracted to /l-/, is deleted in both a and b. The sentence in (a) is supposed to be produced as "muː loːn haːdhə **l-**korsɪ?", whereas the sentence in (b) is supposed to produce as "Fiː foːq **tˤtˤ**aːwleh ħaliːb". The article in sentence (b) is produced as a geminate due to a complete assimilation of the lateral /l/ to the following coronal consonant /tˤ/. Deletion of the article has been attested both in Arabic foreigner talk (e.g. Bizri 2010; Dashti 2013) and Gulf Pidgin Arabic (e.g. Bakir 2010) studies. (9.3) shows the absence of gender agreement, with the masculine form used as the default in most cases. To clarify, the word [sˤufrɪjja] "pot" is feminine in Arabic. Hence, the adjective following it [kabiːr] "big" is supposed to agree with it in gender and should be produced as /kabiːr-ah/. This phenomenon was also attested in foreigner talk (e.g. Dashti 2013) and Gulf Pidgin Arabic (e.g. Al-Ageel 2015) studies.

Findings with regard to the lack of consonant simplification in FDH-directed speech point to a deviation of this speech style from infant-directed speech (IDS). IDS is characterized by modification of consonants for the purpose of simplification as discussed in Chapter 3 (See Ferguson 1964; Ferguson, 1977; Cruttenden 1994). Parents' segmental modification of their speech to their children has been interpreted as a sign of affect and solidarity as suggested by Cruttenden (1994). Comparing this to FDH-directed speech, it is unsurprising that NSs in the current study did not modify their consonants as speech addressed to FDHs is unlike that addressed to infants or children in terms of emotions and affect. Evidence for this comes from FDS studies in which FDS has been more often rated with lower vocal affect compared to IDS or ADS (e.g. Uther *et al.* 2007).

Despite the lack of accommodation in FDH-directed speech in terms of consonant simplification, FDH-directed speech exhibited the use of less marked syllable structures more than ADS, at least with onset clusters (see Chapter 7). However, complex codas did not undergo robust changes. As discussed in Chapter 4, all L1 systems of the FDH participants lacked onset (except Bengali, Swahili, Luganda and Tagalog) and coda clusters. Additionally, production results of complex syllable structure (See Chapter 8) provide evidence of a high tendency by FDHs to produce the less marked syllable structures (CV syllables) word initially and word finally. Stated differently, FDHs' productions were characterized with more retention of the short vowel in unstressed syllables. The tendency to produce less marked onset consonant clusters was significantly higher than the tendency to produce less marked coda consonant clusters. From these results, NSs might have wanted to converge to their interlocutor's interlanguage system by simplifying onset consonant clusters, given that consonant clusters are optional in this position.

It is important to note, however, that the function of adhering to the least marked syllable structure when producing words with consonant clusters might differ based on the context and social factors in the dialect of Omani NSs. For instance, in a sociolinguistic study that investigates the features that

NSs of Nizwa Arabic acquire after migration to the capital Muscat due to their contact with supralocal variants, Ambu Saidi (2019) found a low tendency towards vowel deletion (syncope) among her participants. Females and participants aged between 25-40 were especially likely to retain the short vowel than males or younger participants. She explains this as unsurprising given that women generally, and Arabic-speaking women specifically, have been associated with standard, prestigious forms (e.g. Lynch 2009; Al-Wer 1997; Taqi 2010). As we have seen in Chapter 7, FDH-directed speech and ADS both exhibited divergence from syncope. NSs' linguistic behavior in ADS can be compared to that of Ambu Saidi's participants. Despite the fact that the NS participants in the current study were not migrants like those of Ambu Saidi, high rates of vowel retention in ADS might reflect the same motives as those discussed in Ambu Saidi. Given that the NS adult interlocutor (i.e. the researcher/author of the present thesis) was a stranger and unfamiliar to the NS employers, choosing to address her with a more standard, urban form was not unexpected. Moreover, given that the NS participants in the present study were females, it is possible that they would likely use the more prestigious form based on evidence on females' association with prestigious forms. The NS participants in this study likely had different motives when they chose to use the less marked syllable structure more often when addressing FDHs. NSs would likely not use the less marked syllable structure to sound prestigious or urban when addressing their FDHs as the latter are perceived inferior and foreign. Rather, NSs might have used this syllable structure because it is simpler and would help ease communication.

One might argue, however, that if NSs chose to simplify marked syllable structures to enhance communication, why did not they simplify their marked consonants for the same purpose? Syncope is a phonological process that NSs might or might not apply when they produce initial consonant clusters (Kiparski 2003). Hence, whether to simplify the syllable structure or not becomes an optional matter that is validated in the NSs' standard dialect. Another answer for the previous argument comes from findings on coda consonant clusters. Because neither coda consonant clusters nor consonants undergo variation in Nizwa Arabic, this is a likely reason why NSs did not simplify them even for the benefit of FDHs. NSs preferred to maintain their standard dialect by reinforcing the non-foreign-foreign role differences but not sound foreign. As explained in Chapter 4, breaking up coda consonant clusters is a common phenomenon in other Arabic dialects such as the Levantine dialects (Kiparski 2003). Thus, breaking up the cluster for simplification or comprehensibility purposes would result in NSs of Nizwa Arabic speaking a dialect other than their own.

From a linguistic point of view, most studies agree that in the languages of the world, the CV syllable is universal (e.g.; Vennemann 1988; Clements 1990; Blevins 1995). Historical linguistic studies establish that syllable structure alterations in languages comply with syllable preference rules (Carlisle 2001). For example, Vennemann (1988) has demonstrated that CCV syllables are reduced to CV syllables in some languages. One example comes from Early Old High German, which involved complex onsets consisting of /h/ followed by a consonantal sonorant such as *hnigan* ‘to bow’. In Late Old High German, the /h/ had disappeared resulting in a syllable with one consonant as in *nigan*. Although this requires further investigation, it could be assumed that NSs of Nizwa Arabic abide by syllable preference laws and have generally started showing shifts towards a less marked syllable structure.

8.1.2 Acoustic and prosodic adaptations in FDH-directed speech

Acoustic and prosodic examinations of consonants and vowels in FDH-directed speech were set to test the following hypotheses (H3a and H3d will be discussed in Section 8.1.3):

> **H3.** NSs will hyperarticulate their speech acoustically and prosodically to enhance its clarity for their FDH interlocutors.
>
> **H3b.** Consonant intensity will be higher in FDH-directed speech.
>
> **H3c.** Fricative duration will be longer and VOT of stop consonants will be longer in FDH-directed speech.
>
> **H3e.** Vowel intensity will be higher in FDH-directed speech.
>
> **H3f.** Vowel duration will be longer in FDH-directed speech.
>
> **H3g.** *f*0 will have no change in FDH-directed speech.

H3 was partially supported in the present study. While results on the acoustic examination of consonants did not provide support to H3, results on vowels did. Spectrographic views of the consonants obtained from acoustic profiling showed that the acoustic pattern of all consonants was relatively similar across FDH-directed speech and ADS. However, some of the target consonants showed variation in their acoustic features (e.g. /ʕ/, /ʁ/, /θ/, /ð/ and /ðˤ/) within and across speech style. Evidence was discussed with regard to the nature of variation in the production of these sounds (Chapter 7). Some types of variation in production exhibited by these sounds was in line with similar reports in the Arabic literature, and so cannot be a by-product of who the interlocutor was. Other types of variation especially the stop-like realization of the interdentals proved to occur in both ADS and FDH-directed speech, thus abandoning any doubts on whether it was triggered by who the interlocutor was. A comparison of ADS and FDH-directed speech with regard to the acoustic adaptations of the plain/emphatic fricative consonants revealed a few systematic adaptations along the spectral, temporal and prosodic dimensions of the examined consonants. Results confirmed significantly higher skewness in FDH-directed speech for both /s/ and /sˤ/ and generally higher kurtosis of both /t/ and /tˤ/ but significant changes were confirmed for /tˤ/ only. At the temporal dimension, shorter durations of both fricatives were attested in FDH-directed speech, but significant shortening was demonstrated for /sˤ/. This finding did not provide support to H3c. Additionally, a general pattern of vowel shortening was demonstrated in FDH-directed speech, rejecting H3f. Prosodically, FDH-directed speech exhibited higher intensity in the fricative and stop consonants examined as well as vowels, but among the consonants, a significant rise in intensity was only attested for /tˤ/. This was in line with H3b. Vowel intensity was significantly higher in FDH-directed than in ADS, confirming H3e. One more prosodic difference between FDH-directed speech and ADS was attested in vowel fundamental frequency (F0) which was significantly higher in the former group. This was surprisingly against H3g. These results will be justified in the following paragraphs.

Spectral moments are reported to be reliable acoustic cues in distinguishing place of articulation in both English (Jongman *et al.* 2000) and Arabic (Al-Khairy 2005). Spectral moments, however, failed to distinguish the plain/emphatic fricative consonants investigated in the current study. This is consistent with previous Arabic research that also reported no effect of emphasis on spectral moments. Nonetheless, FDH-directed speech showed a robust increase in skewness of both /s/ and /sˤ/ compared to ADS. To recall, skewness is a statistical measure of spectral tilt and reflects the slop of power distribution (Jongman *et al.* 2000). Both ADS and FDH-directed speech showed negative values of skewness for plain/emphatic fricatives indicating a positive tilt and a predominance of energy in higher frequencies. This is in line with what previous research found about skewness in different fricative places of articulation including alveolar fricatives. For instance, Jongman *et al.* (2000) found that skewness of post-alveolar fricatives was higher than that of alveolar or labiodental fricatives. They suggested that post-alveolar sounds had the strongest concentration of power in the lower frequencies. Similarly, Al-Khairy (2005) found that skewness of the pharyngeals /ʕ/ and /ħ/ was higher than that of the alveolar fricatives /s/ and /z/, with the pharyngeal /ʕ/ having the highest skewness than any other fricative examined. From these studies, we can conclude that sounds produced further back in the oral cavity have more positive skewness and a predominance of energy in lower frequencies. Following these interpretations, higher skewness values of both /s/ and /sˤ/ in FDH-directed speech indicate that these consonants are realized with more negative tilt and concentration of energy in lower frequencies compared to those of ADS. This finding can be better understood in light of Maniwa *et al.'s* (2009) study of clearly produced English fricatives.

Maniwa *et al.* investigate the difference between clear speech and conversational speech along several acoustic dimensions. They found that clear speech diverged from conversational speech in nearly all spectral dimensions they examined especially at central windows. With regard to skewness, Maniwa *et al.* found that clearly spoken fricatives had lower skewness than fricatives in conversational speech. Obviously, FDH-directed speech diverges from that of Maniwa *et al*. To this end, FDH-directed speech cannot be similar to clear speech with regard to spectral characteristics for two main reasons. First, FDH-directed speech did not exhibit any systematic adaptations along the spectral moment metrics examined except for skewness. Second, despite systematic changes in skewness of both /s/ and /sˤ/ in FDH-directed speech, the pattern of change was opposite to that reported by Maniwa *et al.* in the previous account. Whether changes attested in skewness of FDH-directed speech alveolar fricatives suggest that these sounds are produced slightly further back in the vocal tract compared to their counterparts in ADS and whether this means that they are less clear due to their proximity to neighboring fricatives in acoustic space, is not clear and requires further investigation.

Unlike fricatives, M1, M2 and M3 were found to robustly differentiate alveolar plain stops from emphatic stops. Kurtosis (M4) was not found to change significantly due to emphasis. Systematic changes between FDH-directed speech and ADS were only found to occur with regard to kurtosis of /tˤ/. /tˤ/ generally had higher kurtosis than /t/ in both ADS and FDH-directed speech. High kurtosis values indicate high peakedness and a noticeably defined spectrum. This might suggest that /tˤ/ was produced more clearly in FDH-directed speech compared to that in ADS. However, our findings do not point to any other exaggerations in the production of this sound in FDH-directed speech, which leaves this claim subject to scrutiny. Moreover, clear speech studies have rarely investigated spectral moments of stops, which makes comparisons of this finding with previous research very limited.

Temporal adaptations (duration of segments) in FDS are typically justified in relation to speech rate. Slow speech rate has been reported to be a robust characteristic of IDS, clear speech, and foreigner talk (Hatch 1983; Mannon 1986; Derwing 1999; Biersack *et al.* 2005; Kangatharan 2014). Wesche (1994: 226) states that "slower delivery results in more careful articulation, provision of full underlying vowel forms and consonant clusters, and avoidance of contractions, so that word boundaries tend to be more clearly delineated". Furthermore, the H&H theory postulates that careful enunciations of speech result in adopting a hypermode in which vowels and consonants are expected to reach their target form. From this, longer vowels and consonants in speech are taken as an indication of a slower speech delivery in IDS and clear speech research. Vowel duration in speech to FDHs was found to be significantly shorter than that of ADS based on initial analyses. Furthermore, fricative duration was found to be generally shorter in FDH-directed speech, with /sˤ/ showing significant shortening. FDS has only focused on vowel duration and no reports on the duration of consonants have been investigated. Results on vowel duration in FDS research have not been consistent. In fact, more studies have reported no significant difference between FDS and ADS with regard to vowel length and sometimes reduction (e.g. Knoll *et al.* 2015, Knoll and Scharrer 2007; Uther *et al.* 2007; Biersack *et al.* 2005) compared to those that have reported a difference (e.g. Scarborough *et al.* 2007; Kangatharan 2014). Hence, the results of vowel length from the current study support findings from previous FDS studies.

The absence of a difference between FDS and ADS with regard to vowel length in previous FDS studies is striking because if FDS is assumed to serve a similar role to that of IDS or clear speech, then we should expect speech delivery to be slower and thus segments to be longer. Results from the current study are even more surprising because speech to FDHs exhibited significant vowel shortening or reduction. The current study took the initiative to investigate why vowel shortening took place, contrary to expectations. The statistical difference in vowel length between FDH-directed speech and ADS disappeared after a secondary analysis that removed repeated words from data analysis. Thus, exaggerated repetitions of target words in FDH-directed speech was the source for this significant shortening of vowel duration. Yet, the new result regarding vowel length is actually still against common wisdom especially when vowel length is taken as an indication of speech rate. Similarly, the case of fricative shortening in the present study comes contrary to findings from clear speech and IDS. Maniwa *et al.* (2009) reported longer durations of fricatives in clear speech compared to conversational speech. Likewise, Englund and Behne (2006) found Norwegian IDS to exhibit longer durations of /s/ than ADS. Despite discrepant results from past IDS research regarding VOT (e.g. Sundberg and Lacerda 1999; Malsheen 1980), a more recent study by Englund (2005) demonstrated longer VOT values in IDS compared to ADS. Taken altogether, the general pattern of segment shortening in FDH-directed speech and the absence of a difference in segment duration for other sounds (e.g. vowels and plain/emphatic stops) in the present study can be interpreted by referring to the degree of familiarity between the NS female employers and their interlocutors in the task they carried out. NSs in the present study might have not felt a need to slow their speech down to their FDHs having greater familiarity with them compared to the NS adult who was a stranger to most of the female employers. It should be noted also that although phone duration is a robust candidate in speech rate measurement, it is not the only aspect that has been used to measure speech rate. Other possible candidates include the length of the utterance, length and number of words or syllables per second, length and frequency of pauses, etc. (Narayanan and Wang, 2005; Bradlow *et al.* 2003). Hence, the absence of a difference in segment duration in FDH-directed speech does not necessarily indicate lack of evidence for a slower speech rate. The fact that those other candidates were not examined in the present research makes it unclear whether FDH-directed speech is actually slower or comparable to ADS with regard to speech rate. This can be a worthwhile theme for further investigation.

With regard to prosodic adaptations, a rise of sound intensity in speech directed to FDHs can be interpreted as an increase in loudness based on evidence that confirms that intensity is the acoustic correlate of loudness (Fernald and Mazzie 1991). According to Lehiste (1976), a rise in sound intensity is a consequence of an increase in respiratory effort. With other factors kept constant, sounds with higher intensity are perceived as louder (Lehiste 1976). A rise in intensity was attested in both consonants and vowels of FDH-directed speech. The present study has also found that vowels exhibited increased fundamental frequency (*f*0) in FDH-directed speech. A rise in *f*0 is the consequence of an increase in vocal-fold tension (Monsen *et al.* 1978). In combination with other cues including intensity and duration, the adaptations of *f*0 in English emphasize to the listener the syllables that are being stressed in a word or the words that are being emphasized in a sentence (Lehiste 1976). In the production of a stressed syllable, *f*0 is heightened due to an increase in subglottal air pressure and focal-fold tension (Monsen *et al.* 1978). Hence, one interpretation for the heightened intensity and *f* 0 in the target words produced in speech directed to FDHs is an increase in vocal effort. The attested increase in F1 of the vowels examined in FDH-directed speech relative to ADS confirms this interpretation. In a study that examined the effect of vocal effort on vowel spectral properties, Lie´nard and DiBenedetto (1999) reveale that F1 and *f*0 both increase when speakers increase vocal effort. In light of this finding, Ferguson and Kewly-Port (2002) demonstrate that the increase in F1 of vowels produced clearly by the talkers they examined was a consequence of an increase in vocal effort. Additionally, they observed an increase in the intensity of clear speech during recording and digitization, which they attributed to an increase in vocal effort. They concluded that when talkers speak clearly, they might be expected to increase their vocal effort. In the same vein, Picheny *et al.* (1986) further demonstrate the role of increased vocal effort in clear speech when they revealed in a study that clear speech was 5-8 dB more intense relative to conversational speech.

Prosodic alternations in speech to infants have also been confirmed in many studies (e.g. Fernald and Kuhl 1987; Grieser and Kuhl 1988; Cooper and Aslin 1990; Brosesch and Bryant 2015). IDS is characterized with higher pitch (*f*0), expanded pitch contours, higher amplitude and slower temper compared to ADS (Meser 1981; Fernald and Kuhl 1987; Albin and Echols 1996; Adriaans and Swingley 2017). Heightened prosodic characteristics of IDS are claimed to be a major source for infants' preference for motherese (Fernald and Simon 1984; Fernald and Kuhl 1987). Prosodic modifications in IDS have been claimed to offer two main benefits: affect (Fernald and Khul 1987; Knoll *et al*. 2006) and language learning (Segal and Newman 2015; Adriaans and Swingley 2017).

Second language acquisition researchers have come to believe that preverbal infants have a tendency to extract and identify words and syllables within a sentence that are highlighted by some aspects of prosody including pitch and stress (Gleitman and Wanner 1982; Butler *et al.* 2014). This tendency may assist children in identifying boundaries between syllables and words by attending to the highlighted elements since there is no one particular set of acoustic markers for utterance boundaries in languages like English (Albin and Echols 1996). From this, researchers have argued that the exaggerated prosodic cues of motherese can indeed be beneficial for children in solving such a problem. Since FDH-directed speech cannot be similar to IDS in terms of emotions, evidence from clear speech discussed earlier and IDS suggest that changes in the prosody of FDH-directed speech can be best interpreted as serving a didactic role.

The exaggerated prosodic cues in FDH-directed speech do not necessarily serve a similar role to those in IDS with regard to, for example, assisting FDHs in identifying phrase boundaries. This is because the present study did not examine other prosodic cues that are likely to highlight phrase boundaries including pitch range and pitch contours. Moreover, secondary analyses confirmed that phrase boundary did not play a role in prosodic exaggerations evident in FDH-directed speech. However, heightened *f*0 in FDH-directed speech might have been used as an attention-eliciting strategy given that NSs were aware of their FDHs communicative needs.

As explained earlier, English stressed syllables in ADS are distinguished by one or more of the following suprasegmental aspects: higher pitch ($f0$), higher amplitude (louder) and larger duration (Lehiste 1976). Any of these properties (but particularly pitch and duration) can contribute to the perception of stress (Albin and Echols 1996). Variability in intensity is considered the weakest cue to stress while fundamental frequency the strongest (Streeter 1978). Conversely, Arabic studies have demonstrated that Arabic stress is determined by the higher intensity and longer duration of the stressed syllable (Al-Ani 1992; Al-Ani 1991). Additionally, focus words within an utterance tend to receive more stress than others. This emphasis provided by stress tends to play a role in speech processing and word recognition (Cutler and Foss 1977; Grosjean and Gee 1987). In the present study, vowels appeared in stressed syllables in the target words examined. Moreover, the target words were likely candidates for sentential stress (focus) as they represented the objects in the pictures as explained in Chapter 5. Hence, it is very likely that NSs heightened prosodic cues for focus words when addressing their FDHs to highlight them and stimulate auditory attention of FDHs. One argument against this claim is that duration, a major cue for stress, was not exaggerated in FDH-directed speech. Despite the validity of this argument, one cannot deny the fact that target words in FDH-directed speech were produced with an exaggerated vocal effort which leads to louder and more prominent productions as explained earlier. This may assist FDHs in learning these words faster and understanding them in relation to their context.

To this end, findings from phonological, acoustic and prosodic correlates of FDH-directed speech invite two main conclusions. First, FDH-directed speech is not different from ADS at the consonantal level based on auditory and acoustic analyses. This suggests lack of accommodation to either FDHs' language errors or FDHs' foreign accentedness. Second, FDH-directed speech is similar to IDS and clear speech in some of its vocal and prosodic characteristics especially F1, $f0$ and intensity. This indicates that prosodic exaggerations in FDH-directed speech likely serve a didactic role.

8.1.3 Hyperarticulation in FDH-directed speech

Acoustic examination of consonants and vowels was also set to test the following hypotheses:

> **H3a.** NSs will enhance the contrast between plain/emphatic consonants by hyperarticulating the prominent acoustic correlates that the NSs in the current study use to distinguish emphatic from plain consonants.
>
> **H3d.** Vowel space area will be expanded in FDH-directed speech indicating hyperarticulation.

Another possible scenario for the situation I narrated in Section 8.1.1 is that given the perceived foreign accent of the interlocutors, talkers may adjust their productions acoustically in that they might produce clearer instances of their sounds and even enhance the distinctiveness between contrastive sounds in order to improve comprehension or intelligibility. Past literature confirmed that acoustic adjustments of speech to special listeners (e.g. hearing impaired, infants, NNSs) can be made consciously or unconsciously for the benefit of the listener (Pichney *et al.* 1985; Khul *et al.* 1997; Uther *et al.* 2007). Phonetic enhancement of speech sound contrasts (i.e. hyperarticulation) in speech to such listeners is assumed to facilitate language learning or increase intelligibility of speech depending on the needs of the listener as demonstrated by the H&H theory (Lindblom 1990). Phonetic enhancement of stop and fricative sound contrasts is attested in IDS and clear speech research (see sections). Based on this literature, FDH-directed speech was predicted to involve clearer articulations and enhancement or maximization of consonant contrasts that are likely to be confused by the foreign listeners and that have some degree of perceived phonetic difficulty (H3a). Examination of plain/emphatic fricative and stop contrasts, however, did not provide support to this prediction. Interactions between sound type and speech style were not significant, indicating the lack of phonetic hyperarticulation in FDH-directed speech. This finding is against those reported by clear speech and IDS studies about enhancement of consonant contrasts.

In an investigation of clearly spoken English fricatives, Maniwa *et al.* (2009) found that fricative place of articulation contrasts was particularly enhanced by talkers instructed to speak clearly. For instance, nonsibilants are inherently characterized with more diffuse spectra. M2 (standard deviation or variance) increased for non-sibilants and decreased for sibilants. In addition, there was a greater increase in F2 transition distance for palato-alveolars, which inherently have higher F2 compared to that of alveolars. The same change was reported about dentals (with inherently higher F2) compared to labio-dentals. Voicing contrasts among fricatives were also enhanced in clear speech, this showed in how voiceless fricatives, which have inherently longer duration, increased more in duration than voiced fricatives. Similarly, in an investigation of a large corpus of caregivers' speech to their infants and to another adult, Cristia (2010) found that consonantal contrasts were enhanced in IDS. The study examined the sibilant contrast /s/ vs. /ʃ/ and focused on M1 (mean or center of gravity) as the dependent variable. Results from interaction between sound type and register revealed that M1 was significantly increased for /s/ but remained the same for /ʃ/. This was comparable to the finding reported by Maniwa *et al.* (2009) regarding the increase of M1 for /s/ in clear speech but not for /ʃ/.

Findings from consonant contrast analyses in the present study confirm a lack of support to the hyperarticulation hypothesis and a deviation of FDH-directed speech from other speech codes that reported consonantal enhancement, including IDS (but see Benders 2013; McMurray 2013) and clear speech. These findings can be interpreted in relation to methodological and context specific factors. For example, methods used to elicit clear speech are totally different from the method used here to elicit FDH-directed speech or IDS. Because clear speech research is mainly targeted towards developing clear speech training programs that can help individuals with difficulties in understanding speech, it has been elicited using specific cues (Lam *et al.* 2012: 1807). For instance, clear speech has been elicited using different instructional procedures such as "hyperarticulate" (Dromey 2000), or "speak clearly" (Ferguson and Kewley-Port 2007), or "speak to a NNS" (Bradlow *et al.* 2003), or "speak to someone with a hearing-impairment" (Smiljanić and Bradlow 2008), or a combination of these cues. Clear speech studies have suggested that speech output may vary depending on the cue or instruction used to elicit this kind of speech (Lam *et al.* 2012). Additionally, clear speech is elicited in laboratory settings using specific material and highly controlled stimuli. For example, many clear speech studies have used written sentences and/or specific phonetic context to elicit clear speech (e.g. /hVd/, VCV) (e.g. Pitchney *et al.* 1986; Ferguson and Kewley-Port 2007; Maniwa *et al.* 2009; Lam *et al.* 2012). Another way in which clear speech is methodologically different is that speech is challenged by either a background noise or a hearing-impaired listener which makes acoustic hyperarticulation better justified (Smiljanić and Bradlow 2008). The fact that FDH-directed speech was elicited from spontaneous conversations and was not controlled by a specific set of prosodic or phonetic context justify part of this lack of hyperarticulation. Support for this comes from a study by Lam *et al.* (2012) who investigate the effect of different instructions on the elicitation of clear speech. They instructed 12 participants to read 12 sentences in four different conditions: "habitual", "clear", "overenunciate" and "hearing impaired". The results indicated that the four different conditions elicited different amounts of acoustic change. The non-habitual conditions elicited more hyperarticulation of vowel space and vowel spectral dynamic properties compared to the habitual condition. Additionally, maximization was greatest for the “overenunciate” condition compared to the other non-habitual conditions. Although the present study is different from that of Lam *et al.* (2012) in that the speakers in this study were interacting with NNSs in a real interaction setting, it is possible that NSs did not find a need to hyperarticulate their consonant contrasts in the absence of such specific instructions or cues like those of the previous study regardless of the identity of their interlocutors. Additionally, the fact that the NSs had known their FDHs for some time may have not prompted them to hyperarticulate when addressing them.

With regard to the deviation of FDH-directed speech from most studies that characterized IDS as hyperarticulated, context related factors may well justify this finding. Based on an influential study by Khul *et al*. (1997) and subsequent research, caregivers have been found to unconsciously enrich the input for their children and make language acquisition less challenging. The case of FDH-directed speech is different in that the NSs’ might not have felt a need to hyperarticulate specific consonant categories because their FDH interlocutors already know how to speak (though they might have lexical gaps) despite the fact that they have segmental errors. As explained earlier, NSs’ utmost challenge when interacting with FDHs might be communicative more than anything else. Hence, we find them enhancing suprasegmental cues (e.g. loudness and higher pitch) that highlight the specific words they want the FDHs to pay attention to as well as use other communicative adjustments to clarify these words (e.g. repetition, elaboration, etc.). Based on claims of the H&H theory, the goal of speech production is influenced by a number of factors including the social context in which communication takes place (Lindblom 1990). Hence, speakers’ asses the needs of their interlocutors and adapt their speech accordingly.

Despite the potential validity of these interpretations, interestingly, the difference in vowel space area between FDH-directed speech and ADS in the current study proved to be significant, with FDH-directed speech having a slightly larger area than ADS. Expansion of vowel space area has also been widely known to enhance intelligibility as in clear speech (e.g. Ferguson and Kewly-Port 2002) and play a didactic role as in IDS (e.g. Khul *et al.* 1997) and FDS (e.g. Uther *et al.* 2012*;* Uther *et al.* 2007) as discussed in Chapter 3. In clear speech, talkers enhance the difference between vowel contrasts to produce clearer instances of their articulations due to the presence of a communicative barrier (e.g. noise) or speaking to hearing-impaired listeners. In IDS and FDS, when speakers produce vowels farther apart in the acoustic space, they are argued to make the task of learning the new language less challenging for their infant or foreign interlocutors (Kuhl *et al.* 1997; Uther *et al.* 2007). For example, Uther *et al.* (2012) examined whether hyperarticulated speech addressed to native and non-native listeners elicits "larger phonetic change responses", indexed by the mismatch negativity factor of the auditory event. They found that hyperarticulation elicits larger mismatch negativity responses in both groups of listeners, suggesting the benefit of hyperarticulation to listeners. In addition, evidence on vowel hyperarticulation is strongly established in the literature of FDS, clear speech and IDS more than that on consonant enhancement. Thus, my finding on the hyperarticulation of vowels but not consonants might contribute to this research.

8.1.4 The role of foreign accentedness, LoR and religion in FDH-directed speech

Examination of foreign accentedness, LoR and religion in FDH-directed speech was set to test the following hypothesis:

> **H4.** Simplification and hyperarticulation in FDH-directed speech will vary based on FDHs' foreign accentedness score, LoR and religion.

H4 was not supported in the present study. Foreign accentedness score, being Muslims or non-Muslims and LoR did not play any role in whether FDH-directed speech would be hyperarticulated or not contrary to my expectations.

Foreign accentedness of FDHs was assumed to influence the extent of hyperarticulation in FDH-directed speech. One factor that has been found to affect the magnitude of linguistic adaptations in speech to NNSs is foreign accentedness of the NNSs. According to Snow *et al.* (1981), foreigners who are less foreign accented and have more second language production skills may trigger less foreigner talk features. Kangatharan (2014) found that interlocutors who are foreign sounding triggered FDS more than those who were native sounding. Similarly, age or linguistic competence of children has been found to play a role in the linguistic adaptations made by caregivers as seen in Chapter 3. Based on the mother-infant phonetic interaction model (MIPHI), phonetic and prosodic exaggerations in IDS are determined by infants' linguistic development (Sundberg 1998). Findings revealed that consonantal categories are over-specified for children at the word-stage or children beginning to speak but are underspecified for younger infants (Malsheen 1980; Sundberg 1998; Cristia 2010). Based on the literature on foreigner talk and IDS, the present study predicted that the extent of adaptations in FDH-directed speech will be determined in part by FDHs' foreign accentedness. Results, however, showed that foreign accent rating did not trigger hyperarticulation in FDH-directed speech. This might be due to evidence from Chapter 6 that revealed that FDHs' foreign accented rating was very similar. The highest percentage of rating to the FDH group went to the "not at all native-like' ranking on the scale. Thus, NSs might have perceived all FDHs as foreign accented. Therefore, the extent to which hyperarticulation was achieved or not did not depend on the degree of foreign accent.

LoR was presumed to correlate with the Arabic experience FDHs had. Hence, the longer FDHs had stayed in the Arabic-Speaking World, the more experience of Arabic they were assumed to have gained. LOR did not influence the degree of hyperarticulation in the present study. This finding can be interpreted based on the peculiar FDH context. The perception and production experiment that the FDHs carried out in the present study showed that LoR had no influence on the learners' performance (see Chapter 8 for results and section 9.2 for an elaborated discussion). Other factors played a role in FDHs L2 performance such as their L1 and L2 literacy as well as their L1s. Moreover, FDHs could vary greatly with regard to the quantity and quality of input they receive in the target community, which might affect their experiences. This suggests that LoR might not be correlated with the learners' Arabic language experience after all. Also, Muslim FDHs were predicted to trigger less hyperarticulation compared to non-Muslim FDHs. This was due to the fact that NSs were Muslims and might be expected to be more familiar with the Muslim FDHs' Arabic language skills. However, religion did not determine the extent of hyperarticulation in this study.

To this end, this study confirms that FDHs characteristics did not determine whether hyperarticulation will be applied or not.

9.2 Interpretation and Implications of FDHs' Performance in L2 Perception and Production

Examination of FDHs' L2 perception and production was proposed to answer the following research questions:

Q2) To what extent are FDHs perceptually sensitive to the phonemic contrasts of their L2 Arabic? What social and cognitive factors (independent variables) are responsible for any variation present in the listeners' performance?

Q3) How accurate (or native-like) are FDHs at producing Arabic consonants and complex syllable structures? What social and cognitive factors (independent variables) are responsible for any variation in the accuracy level of domestic helpers' production skills?

Hypotheses set to test the first part of Q2 and Q3 are:

H5. FDHs' perceptual sensitivity towards Arabic phonemic contrasts will not reach native-like performance.

H7. FDH will fall short of target-like accuracy in producing Arabic consonants.

H9. FDH will fall short of target-like accuracy in producing Arabic consonant clusters.

H9.a FDHs will be more successful at producing onset consonant clusters than coda consonant clusters due to the markedness of the latter.

A general finding that this study can confirm with regard to FDHs' sensitivity to consonant contrasts in the present study is their lower performance compared to that of the NSs. However, FDHs' were successful at discriminating some of the *different* pairs. With regard to their production of complex Arabic consonants and consonant clusters, a major finding is that FDHs' spoke with a foreign accent regardless of their LoR and had not yet fully acquired the L2 consonants or consonant clusters. These results support H5 and H7. The rate of acquisition of coda consonant clusters was higher than that of onset consonant clusters. This does not support H9a. Different cognitive and social factors were found to contribute to the participants' performance (Section 9.2.1). Interpretations and implications of these findings are discussed below.

One primary interpretation for FDHs' non-native like L2 performance whether in L2 perception or production is their age of arrival (AoA) to the L2 community. It should be noted that when I use the term AoA, I also mean age of significant exposure or age of onset (AO) as referred to in the literature (Muñoz 2008). This is different from age of learning (AoL) which denotes their first encounter with Arabic. The FDH participants had a mean AoA of 27 years which indicates that their age of significant exposure to L2 input in the target community had passed the critical period for language acquisition. Hence, from a theoretical point of view, as the proficiency levels FDHs achieved fell very short of nativelike proficiency, this provides support to the CPH (Lenneberg 1967) and generally to maturational- and neurological-based interpretations (Scovel 1988; Patkowski 1989; Khul 2007). A plethora of studies on age effects in SLA has provided evidence for the notion "the younger the better". For example, A study by Flege *et al.* (1999) report that Korean immigrants to the United States had stronger accents and lower scores in grammaticality judgment tests than those who arrived earlier in the United States. The researchers attributed these differences to age-effects and the existence of a critical period. It is important to note that discussions of the critical period do not usually differentiate between production and perception (Flege 2003). Stated differently, both perception and production of L2 speech are constrained by a critical period. However, L2 acquisition might affect perception and production differently as discussed in Chapter 2 (Scovel 1988; Flege 2003). It was not the intent of this research to investigate the relationship between production and perception and thus the extent to which segmental perception and production are different with regard to the critical period can be verified by future research. While a late AoA is a valid justification for FDHs' non-native like performance in the current study, it is essential to note that some FDHs learned the Arabic alphabet at a younger age via reading the Qur'an. Hence, interpretations of the CPH can only apply when other factors such as the effect of L2 literacy[25] are taken into consideration (the effect of L2 literacy will be discussed later).

An alternative interpretation for FDHs L2 performance is an influence from their L1 system. In Chapter 4, we have seen that the L1 phonological inventories of all FDH participants lacked the complex Arabic consonants investigated in the current study. Additionally, FDHs' native phonological systems lacked onset and coda consonant clusters. Current theoretical models such PAM by Best and colleagues (e.g. Best *et al.* 2001; Best 1995) and SLM by Flege and colleagues (1988; 1995; 1999; 2002) claim that difficulty in mastering a native-like proficiency in the L2 is not constrained by maturational constraints but rather arise as the result of interference from prior native phonetic learning. Although the main focus of these models is on the perception of speech, SLM has implications for speech production as well (e.g. Flege 1987). The current study was not designed to test any of the claims of the previous models due to the diversity of the FDHs' L1 backgrounds and difficulty conducting cross-linguistic comparisons of their sounds' phonetic and articulatory features. However, FDHs' success at discriminating some different pairs presumably supports Flege's (2002) claim that the mechanisms infants use to learn their native language remain intact across the life span and continue to be accessible for use in L2 learning. Another contribution this study can make in this realm is the evidence it provides for a potential effect of the L1 sound systems of the FDHs in perceiving and producing the Arabic sounds. Interference from the L1 or transfer is most likely the basis for the different foreign accents of individuals from different L1 backgrounds (Flege and Davidian 1984). Flege (1981) claims that the most readily obvious basis for a foreign accent is mispronunciation errors that lead to the perception of a sound substitution like in French-accented English *I sink so* for *I think so*. A popular example in the literature with this regard is Japanese

[25] This is because, presumably the age at which these foreigners started learning Arabic was the same when they started school.

English speakers' productions of English /l/-/r/ as Japanese /r/ which makes these speakers easily identified by their accent (Aoyama *et al.* 2004). Similarly, the finding that the FDHs in the current study were foreign accented was drawn from the perception of segmental substitutions in their production of Arabic complex consonants (See Figure 8.7 in Chapter 8 for evidence).

The two interpretations provided so far for the non-target like the performance of the FDHs belong to two major proposals in the SLA literature to explain L2 learning. Despite their significant contribution in explaining the differential learnability of the L2, the present study isolated these two factors from inclusion in the statistical models used in examining the factors that affect perception and production of L2 Arabic. This is because it was assumed that the FDHs were comparable in terms of their AoA and L1 background (except with regard to initial consonant clusters, which were analyzed qualitatively). The FDHs, indeed, migrated to the Arabic-speaking Middle East as adults and so their age of significant exposure to Arabic was comparable. L1 backgrounds of the FDHs all lacked the complex Arabic consonants and final consonant clusters examined in this study and therefore were assumed comparable in this regard. To account for individual variation in FDHs' L2 performance, other factors were considered including the L1 (with regard to onset consonant clusters only), LoR, L1 schooling and L2 literacy. Findings on these will be discussed in detail in Section 9.2.1).

But first, it is essential to shed some light on variation in FDHs' performance with regard to onset and coda consonant clusters. The results revealed that the difference in the rate of modification between onset and coda consonant clusters was significant, with a considerably higher rate of modification attested in coda consonant clusters compared to onset consonant clusters. Stated differently, FDHs produced consonant clusters word finally more often than they did word initially. Theoretically, this is against the markedness principle, which states that final consonant clusters are less frequent in the world languages and thus are more marked (Greenberg 1978; Anderson 1987). Input the FDHs are exposed to can enlighten us in this regard. In light of the findings we obtained on the frequency of onset and coda consonant clusters in FDH-directed speech, NSs' productions nearly contained no modification in coda consonant clusters. Coda consonant clusters were maintained in the speech of the NSs and did not undergo any simplification. However, despite the phonological rule of syncope Nizwa Arabic is known for, and which results in consonant clusters in unstressed open syllables, NSs' productions were characterized with fewer onset consonant clusters and more retention of the vowel in both FDH-directed speech and ADS. This tendency was significantly larger in FDH-directed speech relative to ADS. Hence, due to the optionality of onset consonant clusters and the fact that onset consonant clusters were less frequent in the input, FDHs acquired the less marked structure that was readily available in the target language input. On the contrary, because coda consonant clusters were frequently heard by FDHs in the input they receive, their productions contained more instances of these. This also demonstrates the learnability of this syllable structure, though some FDHs' productions still exhibit modifications of the clusters.

9.2.1 The role of the social and cognitive factors

Examination of the social and cognitive factors that affect L2 speech perception and production was proposed to test the following hypotheses:

> **H6.** FDHs with longer LoR, who have more formal schooling and who are literate in their L2 will be more sensitive to Arabic phonemic contrasts.
>
> **H8.** FDHs with longer LoR, and who have more formal schooling and who are literate in their L2 will be more accurate at producing consonants and consonant clusters.
>
> **H10.** FDHs with longer LoR, and who have more formal schooling and who are literate in their L2 will be more successful at producing Arabic consonant clusters.

9.2.1.1 LoR

The duration FDHs had spent residing in the Arab world appeared to play no role neither in their L2 perception of consonant contrasts nor in their production of Arabic consonants or consonant clusters, rejecting part of H6, H8 and H10. In some cases, FDHs' scores correlated negatively with LoR, as in the production of coda consonant clusters. These findings are against the assumption that the more L2 input one receives the better the opportunities for the NNS to master the L2 (McAllister 2001). Findings on the role LoR plays in NNs' L2 performance are inconsistent in the SLA literature. Findings from the present study are in line with research that did not find an effect for LoR on L2 learners' performance (e.g. Oyama 1976; Flege 1988, 1993).

It is hard to ignore the different ways in which a lack of effect of LoR has been interpreted in the literature. For example, Flege and Liu (2001) suggest three major interpretations discussed in Chapter 2: (1) the amount of L2 input is not a crucial predictor of L2 performance, (2) L2 performance is constrained by a critical or sensitive period, (3) LoR provides a meaningful index of

L2 input for some individuals but not others. In support of the first interpretation, Moyer (2009) argues that LoR is not a reliable determinant of L2 phonological proficiency; i.e. mere exposure in the L2 is not sufficient for late L2 learners. In order for late learners to gain sufficient input, they need to engage in the L2 environment in different ways. Situations favorable to such attempts vary across individuals depending on age as well as educational, social and ethnic background. This is why different individuals in different settings obtain various scores, which supports Flege and Liu's (2001) third interpretation. Moyer (2009) and Moyer (2004) suggest that for LoR to reliably index L2 experience, an integrated approach that takes into account cognitive, psychological and social factors needs to be carried out. An example that illustrates this comes from a study by Moyer (2004). The oral performance of immigrants to Germany with extensive experience in the language was examined. Moyer found that immigrants' performance on a number of tasks was correlated with particular L2 input predictors such as instruction in the L2, contact with native speakers, sense of self identity in the L2, contexts of L2 use and others. Findings from this study confirmed the following conclusions. Psychologically, LoR correlates to a sense of overall fluency and satisfaction with L2 attainment as well as motivation to learning the L2. Socially, LoR correlates with the frequency of contact with NSs and intention for permanent residency in the L2 target community. Cognitively, LoR correlates with L2 instruction and communicative use of the L2 rather than just focusing on form as well as the amount of feedback on pronunciation and kind of phonological training. Another study by Flege (2002) found that LoR can play a role in adults' L2 performance only if they are exposed to a considerable amount of native-speaker input. In this study, Chinese students with longer LoR in the United States were significantly better in their L2 performance compared to students who had shorter LoR. LoR, however, did not predict the performance of non-students. Flege (2002) concludes that because students had more opportunities for receiving NS input compared to the non-students, their performance improved noticeably over time.

In light of the previous observations and findings, it is not surprising that LoR was not found to play a role in FDHs' phonological performance. In order to fully understand the reason behind the lack of effect of LoR, it is essential to investigate the FDH-context in light of Flege's interpretations and Moyer's (2004) integrated model. Despite the fact that FDHs are exposed to NS input in a naturalistic setting, and despite evidence provided by the current study on the absence of any simplification of the consonants of interest, this exposure does not guarantee the amount or quality of input they are getting. In fact, FDHs' language contact with their employers or with other family members can be restricted to conversations around home chores. In addition, the input they receive can be varied as they can attend to children who have not yet fully developed their native-language sounds and can hear accented-Arabic from other FDHs. Flege and Liu (2001) state that immigrants to North America are likely to use their L2 English with other NNSs as a lingua franca which results in inauthentic L2 input. Additionally, FDHs' employment is temporary and therefore they are aware that their residency in the target language community is not permanent (Bizri 2014). Moreover, FDHs are exposed to L2 input in a naturalistic setting compared to an instructional setting. Thus, they do not get any kind of training on L2 phonology or other linguistic aspects. Additionally, they do not get any feedback on their pronunciation as I confirmed in Chapter 1. In a study by Flege *et al.* (1999), the scores of 240 Korean adults on a grammaticality judgment test were found to significantly correlate with years of education in the United States. All in all, opportunities for meaningful input, contact with native speakers and L2 instruction are very limited in this context. This supports the assumption that LoR is not a reliable index of L2 experience (Moyer 2009; Moyer 2004; Flege and Liu 2001). To reliably examine to which extent FDHs' L2 proficiency level affects their L2 phonological performance, future research needs to take into account how cognitive and socio-psychological factors correlate with LoR first.

9.3.1.2 Literacy

Although LoR was not a significant predictor of FDHs' phonological performance, L1 schooling and L2 literacy appeared to play a role in some of the variables examined. L1 schooling correlated significantly with FDHs' perceptual sensitivity scores and rate of final consonant cluster productions. However, no effect of L1 schooling was found in the production of Arabic consonants or initial consonant clusters. L2 literacy generally correlated positively with sensitivity scores and production of Arabic consonants as well as initial and final consonant clusters. FDHs who reported having learned Arabic via recitation of the Qur'an performed better than those who reported having no literacy skills in Arabic. However, the relationship between L2 literacy and L2 phonological performance was only significant with regard to the production of L2 consonants.

The acquisition of literacy imposes consistent change on youngsters' conception of language (Bialystok and Hakuta 1999). We have learned in Chapter 2 that children who learn the alphabetic script develop a higher level of phonological awareness than others who do not (Morais *et al.* 1979; Adrian *et al.* 1995; Tarone and Bigelow 2005). Despite the scarcity of research that investigates the effect of literacy on L2 phonological performance, literacy is a likely candidate to influence L2 acquisition, particularly in situations of immigration that involve populations that vary greatly with regard to their L1 and L2 literacy (Bialystok and Hakuta 1999). Results on the effect of L1 formal schooling and L2 literacy on FDHs' performance will be discussed separately below.

- **L1 Schooling**

The positive correlation between L1 schooling and L2 performance was not attested in all skills examined. This might be due to the nature of the skill and whether it requires complementary skills learned in school settings to achieve it. For example, the amount of schooling was significantly correlated with FDHs' sensitivity to phonemic contrasts and production of final consonant clusters but not with their production of complex Arabic consonants or initial consonant clusters. Little is known of whether L1 literacy affects perception or phonological sensitivity of L2 consonant contrasts. Thus, it is hard to compare findings from this study with findings from research that investigated the role of L1 or L2 literacy on oral language processing. The latter was concerned primarily with examining the effect of knowledge of the alphabetic script on phonemic awareness and reading. According to Strange and Shafer (2008: 159), perception is "an internal mental (and physiological) process by which the perceiver recognizes incoming stimulus events as instances of mental categories". Similarly, Adrian *et al.* (1995) claims that phonological sensitivity is a universal competence which is why literate and non-literate adults did not differ in their phonological sensitivity in their study. Also, most studies that investigate adult L2 discrimination ability of L2 contrasts rely mostly on interpretations pertaining to the role of the native language phonology or experience on the L2 phonology. To this end, it is unclear whether FDHs' higher scores in L2 discriminability was due to their phonemic awareness and access to an alphabetic script or a consequence of something else. However, a possible justification for this finding is that the AX discrimination task was an experimental paradigm that required the listeners to understand task instruction and procedure as well as sit for an actual test. This protocol might be more familiar to adults who had attended school and experienced such a situation compared to adults who had little to no experience with carrying out cognitively demanding tasks due to not attending school. Certainly, the more schooling a FDH had the more used to taking tests and listening to instructions she would be. These factors might have improved the performance of the more educated FDHs' in the AX discrimination task.

Production of L2 consonants, on the other hand, is a complicated motor control activity that requires the activation of certain articulators in the production of each consonant (Guenther 1994) as well as the establishment of categories for the new features (Flege 1999). Hence, a speaker's job when asked to name pictures, for example, is to remember the name of the object and produce it. The speaker's accuracy in production will then highly depend on their experience with these sounds in the target input and the extent to which they have established categories for the new sounds. That is why, in my opinion, FDHs' L2 literacy played a more significant role here than the amount of their L1 schooling. FDHs who learned Arabic via recitation of the Qur'an must have established categories for the new sounds and gained more experience with complex Arabic sounds compared to those who did not have access to Arabic before moving to the Arabic-speaking World.

Production of consonant clusters requires 'noticing' of structures in the target input and adjusting L1 phonology accordingly. Schmidt's Noticing Hypothesis (1990) claims that conscious awareness (i.e. noticing) of the input plays a substantial role in the process of language acquisition. Several researchers have provided support to this hypothesis and confirmed the importance of noticing for language learning (e.g. Gass 1988; Skehan 1998; Lynch 2001). Among the factors that Schmidt claims to affect noticing is frequency. When a structure or an item appears more frequently in the input due to recurrent instruction or by way of teacher talk, the probability that this structure or item will be incorporated into the interlanguage increases (Cross 2002). The frequency of consonant clusters in final word position was much higher than that of consonant clusters word initially in FDH-directed speech as explained earlier. Hence, the likelihood of FDHs' noticing of the former structure was higher. The significant role the amount of L1 schooling plays in developing conscious phonological processing might have helped improve literate learners 'noticing' skills that are required for L2 learning. Probably this is why the more educated the FDHs were the more successful they were at producing final consonant clusters. The lack of such findings with regard to onset consonant clusters might be due to the optionality of these clusters in NSs speech and their less frequent appearance in the input.

- **L2 Literacy**

FDHs' who were literate in Arabic had more advantages in perceiving and producing Arabic consonants and consonant clusters than those who were not. This is in line with Arabic studies that found advanced learners of Arabic better than those with lower proficiency or non-learners in perceiving and producing Arabic phonemes (See Chapter 2). The more exposure to the target language one has via reading or writing the better their performance is. The significant difference between literate and non-literate learners with regard to the production of complex Arabic consonants particularly might be due to the high-demanding articulation and motor skills these sounds require to be produced. For the literate leaners, learning to read the Qura'n was probably the main task they were required to do during their exposure to Arabic in their home countries. As mentioned in chapter 5, learning the Qur'an fundamentally involves learning the Arabic alphabet and the properties of sound articulations (Supriyadi and Julia 2019). Thus, literate speakers must have acquired properties of Arabic sound articulation since they were children and maintained this mastery as they kept reading the Qur'an in their daily life. On the contrary, there was no statistical difference between literate and non-literate learners in discriminating consonant contrasts and producing consonant clusters. To this end, these findings support Scovel's (1988) claim that compared to all other aspects of language, pronunciation is a physical activity and requires a strong talent for using the articulators.

9.3.1.3 The L1

Before discussing FDHs' performance on initial consonant clusters relative to their L1, it should be noted that FDHs' productions generally contained a high amount of modification of initial syllable structure. This supports Eckman's claim that learners will acquire less difficult L2 structures before acquiring more difficult structures (Eckman 1985). Among the two optional initial syllable structures in Arabic, the CV pattern is less difficult than the CCV pattern because it is more common among world languages (Rice 2007; Zec 2007). Another potential explanation for the speakers' high tendency to produce the CV rather than the CCV pattern is the input they receive from NSs as explained in Section 9.3. As we have seen in Chapter 6, NSs' productions contained higher amounts of the CV syllable word initially compared to the CCV syllable although the common pattern used by the speakers of the dialect is that of the CCV (syncope). FDH-directed speech contained a significantly higher number of the former pattern compared to ADS. Thus, FDHs' productions reflected what they more frequently hear in the target language input.

The L1s of the speakers were considered a potential source for variation in the production of onset consonant clusters due to some differences in whether speakers' L1 systems involved onset consonant clusters or not. Most speakers whose L1 systems lacked onset consonant contrasts including Oromo, Sinhala, and Telugu found it hard to produce the marked syllable structure word initially. This indicates that these speakers had not yet mastered variation existing in the L2 with regard to the initial syllable structure. Specifically, they had not acquired the marked syllable structure consisting of two-member consonant clusters. Although the Yoruba and Indonesian Arabic speakers' L1 systems also lacked initial consonant clusters, they showed some acquisition of initial consonant clusters as their productions contained some instances of consonant clusters. All speakers whose L1 sound systems contained consonant clusters showed variation in their production of onset syllable structure. Their production sometimes involved consonant clusters and other times an epenthetic vowel was inserted between the clusters. The Contrastive Analysis hypothesis (CAH), developed by Lado (1957), can to some extent explain these individual differences. The difference between speakers who did not produce consonant clusters and those who did depend largely on similarity or differences between their L1 syllable structure and that of the L2 (excluding the Yoruba and Indonesian Arabic speakers). CAH postulates that leaners will face difficulty with the L2 when the L1 and the L2 differ. This supports our findings. Similarly, the Markedness Differential hypothesis (MDH) claims that L2 learners will face difficulty with areas of the L2 that are different from the L1 or relatively more marked (Eckman 1985). Undoubtedly, a two-member consonant cluster is more marked than a CV syllable (Anderson 1987). Best (1992), likewise, states that the production patterns of L2 speakers are strongly shaped by their native language phonology. These well-grounded theoretical views provide evidence to our finding that FDHs' L1 shaped their L2 production of complex syllable structure.

Chapter 10. Conclusion and Future Directions

10.0 Introduction

This Chapter will first present the limitations of the present study and the future directions. It will finally provide summary and concluding remarks to the entire thesis.

10.1 Limitations and Future Directions

As with any study of this nature, the results need to be considered in light of unavoidable limitations. First, the diversity of the FDH sample might have affected the findings of this research. For example, FDHs varied on aspects of the study design and which might have affected their abilities in Arabic. For instance, some FDHs had been exposed to Arabic via Islam since childhood. Hence, the diversity of this group might have been the reason behind the lack of effect of LoR, whether in FDH-directed speech or L2 phonological learning. The selection of this sample for the study was due to time constraints and difficulty getting access to a well-defined group in terms of linguistic and social background characteristics. As explained in Chapter 5, FDHs are a vulnerable group and getting families to agree to meet with them was not easy. Hence, an interesting direction for future research is to replicate the current study with FDHs who are comparable with regards to their L1, native language literacy and L2 literacy to examine whether LoR will trigger variable features of FDH-directed speech and whether it will play a role in FDHs L2 performance.

Second, given that the task used to elicit FDH-directed speech lent itself to natural interactions, this made it hard to control for the specific structures and prosodic positions of the target words. Due to this, influences from prosody or neighboring linguistic segments might have affected the results. Yet, the current study accounted for the influence of prosodic position and excluded tokens that were likely to influence acoustic measurements where possible (See Section 6.4.1 on data cleaning). I was, also, constrained with the target words I could include in spot the difference task for two main reasons. First, I was cautious not to include target words that would be unfamiliar to FDHs or would be hard for NSs to describe. Second, it was sometimes hard to find keywords that relate to a specific theme and that start with one of the target consonants chosen for the study. For future research to avoid potential consequences of natural speech elicitation tasks, another task that is more controlled can be used in addition to the more natural task used in this study. It would be interesting to compare results from both tasks and find out whether the more controlled task would trigger features more similar to those reported in IDS, clear speech or FDS.

Fourth, evidence on lack of hyperarticulation with regard to consonants and slower speech in FDH-directed speech needs to be verified by examining other variables. The current study examined only a couple of contrastive consonants. This was due to the time and scope constraints of this study. Future research may examine other contrasts available in the language. In addition, whether FDH-directed speech is not slower than ADS needs to be verified by examining other cues for speech rate such as number of syllables per minute and number and length of pauses. The current study provided an interesting finding with regard to the prosodic features of FDH-directed speech. These findings were interpreted in terms of similar findings in IDS. However, future research can expand the variables used to examine prosodic features and look at pitch contours and pitch range. Such variables can shed more light on the role of prosodic features in FDH-directed speech. Also, bearing in mind that the experimenter was the comparison group might have biased the way NS' modified their speech to either interlocutor group. The lack or very small hyperarticulation effect might be because NSs' were more familiar with their FDHs than the experimenter who was a stranger to them. Future studies need to take this into consideration when designing their studies.

Finally, the AX-forced choice task might not be as accurate as other more complex tasks (e.g. AXB or ABX) in measuring discrimination of consonant contrasts. In addition, this task may be biased as subjects may only answer 'different' if they are not very certain of their response (Gerrits and Schouten 2004). However, the use of the AX task in the current study is well justified. The FDHs are limited in their native education and L2 communicative abilities. Hence, testing their performance using more complex tasks would not have been possible given the complexity of instructions and high memory load such tasks require (Gerrits and Schouten 2004; Strange and Shafer 2008). Besides, most of these tasks have been designed for use with advanced, educated learners in modern societies. Bigelow and Taron (2004) postulate that one of the challenges for studying the effect of literacy on NNSs' awareness and noticing of specific forms is that methods for directing learners' attention require high reading and writing demands. They suggested that most of SLA research on noticing cannot be replicated with illiterate learners. Future research is urged to design methods suited for low-educated L2 learners.

9.2 Conclusion

This study took the initiative to investigate phonological and acoustic properties of speech addressed to female foreign workers exposed to naturalistic input as well as examine the effect of non-linguistic factors such as education and literacy in their perception and production of Arabic consonants and consonant clusters. It also had the advantage of investigating these aspects in a language other than English, which dominates SLA research. Other than being less involved in research, Arabic has a unique set of consonants (the emphatics and the gutturals) that proved difficult for second language learners to acquire, which gives researchers examining these consonants the potential to contribute to the wider SLA research. The study was designed to investigate whether linguistic adaptations that lead to accommodation or hyperarticulation in speech to FDHs occur by examining the phonological and acoustic features of input in a unique NS-NNS interaction setting, that of female employers and their foreign domestic helpers (FDHs). It was also designed to examine the extent to which FDHs' discriminability of Arabic consonant contrasts and their production accuracy of complex L2 consonants and consonant clusters is native-like. The role of non-linguistic factors (e.g. LoR, L1 and L2 literacy) in L2 perception and production were also investigated.

Findings from the present study demonstrate the distinctiveness of FDH-directed speech. FDH-directed speech stands out as a speech style from other speech registers such as FDS, IDS and clear speech with regard to the phonological and acoustic features it exhibits. In this study, the FDH-directed speech context was critical in determining the types of accommodation and lack or small hyperarticulation attested in the speech of the female native speaker employers. These findings urge us to interpret the use or non-use of a speech style in relation to both the speaker and the interlocutor and consider social factors as key in this relation (Lindblom 1990). For example, lack of phonological and acoustic adaptations in FDH-directed speech can only be understood by referring to the FDH context and considering the NSs' perception of these foreigners. At the same time, exaggeration of prosodic features and hyperarticulation of vowel space in FDH-directed speech point to how NSs' perceive the linguistic needs of this group of foreigners. Apparently, cues that would highlight focus words and would elicit the foreigners' attention were considered more important by the NSs to achieve a successful communication. To this end, my findings highlight the contribution of two theoretical models in interpreting FDH-directed speech: The Communication Accommodation theory and the H&H theory.

Results from the present study also shed light on the perception and production ability of a group of uninstructed low-educated foreigners acquiring the language in a naturalistic setting. The status quo in SLA research has been to investigate the cross-linguistic performance of highly educated adults and focus on LoR and AoA as primary factors affecting these learners' performance. Results from this study highlight the importance of looking at low-educated learners and investigating other non-linguistic factors such as L1 schooling. Perception scores of learners in this study were not found to be influenced by their LoR or L2 literacy but the amount of their L1 schooling. Additionally, this study implies that in real life, subjects who are strictly defined based on their L1, LoR, quantity, and quality of exposure to L2 are hard to find. Results from this study also show that whether a factor plays a role in L2 performance depends on the L2 skill examined. For instance, the extent to which L1 schooling and L2 literacy affected L2 performance depended on whether the skill examined was the perception of consonant contrasts, production of consonants or consonant clusters. Results also contribute to the wider SLA research in providing support to research on age effects and L1 interference in SLA. The non-native like performance of FDHs is attributed mainly to their age of extensive exposure to the target language which had passed the critical period as well as to influence from their L1s.

References

Abu-Haidar, F., 1996. 'The Arabic Dialect of Qifṭ (Upper Egypt)'. *Journal of the Royal Asietic Society 6*(1): 89-90.

Adrián, J., J. Alegria and J. Morais. 1995. 'Metaphonological abilities of Spanish illiterate adults'. *International Journal of Psychology 30*(3): 329-351.

Adriaans, F. and D. Swingley. 2017. 'Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability'. *The Journal of the Acoustical Society of America 141*(5): 3070-3078.

Akahane-Yamada, R. 1995. 'Age and acquisition of second language speech sounds perception of American English/r/and/l/by native speakers of Japanese'. *Speech Perception and linguistic Experience*.

Al-Ageel, H. 2015. 'Requests in Saudi Pidgin Arabic'. *Business Management and Strategy 6*(1): 111-137.

Al-Ani, S. 1970. *Arabic Phonology: An Acoustical and Physiological Investigation.* Walter de Gruyter.

Al-Ani, S. 1991. '*Lexical Stress Variation in Arabic: An Acoustic Spectrographic Analysis'*. In *Proceedings of the Colloquium on Arabic Grammar. Budapest*. 9-27.

Al-Ani, S. 1992. 'Stress variation of the construct phrase in Arabic: A spectrographic analysis'. *Anthropological Linguistics*: 256-276.

Albin, D. and C. Echols. 1996. 'Stressed and word-final syllables in infant-directed speech'. *Infant Behavior and Development 19*(4): 401-418.

Alfallaj, F. 2016. 'Foreigner talk and communication strategies: A socio-linguistic study of interactions with foreigners in Saudi Arabia'. *Theory and Practice in Language Studies 6*(1): 40-45.

Ali, L. and R. Daniloff. 1972. 'A contrastive cinefluorographic investigation of the articulation of emphatic-non-emphatic cognate consonants'. *Studia Linguistica* 26: 81-105.

Alghamdi, E. 2014. 'Gulf Pidgin Arabic: A Descriptive and Statistical Analysis of Stability.' *International Journal of Linguistics 6*(6):110.

Algryani, A. 2014. 'Emphasis Spread in Libyan Arabic'. *International Journal on Studies in English Language and Literature 2*: 30-38.

Al-Khairy, M. 2005. 'Acoustic characteristics of Arabic fricatives'. Unpublished Doctoral Dissertation, University of Florida.

Al-Karouri, A. 1996. *Phonetics of Classical Arabic: A Selectional Study of the Problematic Sounds*. KUP: Khartoum University Press.

Al Khatib, S. 2008. 'On the directionality of emphasis spread'. In *Proceedings of the13th Annual Conference of the International Speech Communication Association.* 1-10.

Al Mahmoud, M. 2013. 'Discrimination of Arabic contrasts by American learners'. *Studies in Second Language Learning and Teaching 3*(2): 261-292.

Al-Masri, M. and A. Jongman. 2004. 'Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results'. In *Proceedings of the 2003 Texas Linguistics Society Conference. Somerville, MA: Cascadilla Proceedings Project*. 96-106.

Almoaily, M. 2014. 'Language Variation in Gulf Pidgin Arabic.' In I. Buchstaller, A. Holmberg and M. Almoaily (eds) *Pidgins and Creoles Beyond Africa-Europe Encounters*. Amsterdam: John Benjamins Publishing Company. 57-83.

Al-Najjar, S. 2004. 'Women migrant domestic workers in Bahrain'. In S. Esim and M. Smith (eds) *Gender and Migration in Arab States: The Case of Domestic Workers*. Regional Office for Arab States. Beirut. 25-41.

Al-Nuzaili, A., 1993. '*Experimental study of emphasis and voicing in the plosives of Yemeni Spoken Arabic with some implications for foreign language teaching and learning'*. Unpublished Ph.D Dissertation, University of Leeds.

Alosh, M. 1987. 'The perception and acquisition of pharyngealized fricatives by American learners of Arabic and implications for teaching Arabic phonology'. Unpublished Ph.D Dissertation, The Ohio State University.

Alotaibi, Y. and G. Muhammad. 2010. 'Study on pharyngeal and uvular consonants in foreign accented Arabic for ASR'. *Computer Speech & Language 24*(2): 219-231.

Al Suwaiyan, L. 2018. 'Diglossia in the Arabic Language'. *International Journal of Language and Linguistics 5*(3): 228-238.

Al-Tamimi, J. and M. Barkat-Defradas. 2003. 'Inter-dialectal and inter-individual variability in production and perception: A preliminary study in Jordanian and Moroccan Arabic'. *In 5th AIDA Proceedings.* Associatio Internationale de Dialectologie Arabe. 171-186.

Al-Tamimi, J. 2017. 'Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations'. *Laboratory Phonology: Journal of the Association for Laboratory Phonology 8*(1): 1-40.

Alwabari, S. 2013. *'Non-native production of Arabic pharyngeal and pharyngealized consonants'*. Unpublished M.A Dissertation, Carleton University.

Alwan, A. 1989. ‘Perceptual cues for place of articulation for the voiced pharyngeal and uvular consonants’. *The Journal of the Acoustical Society of America 86*(2): 549-556.

Al-Wer, E. 1997. ‘Arabic between reality and ideology’. *International Journal of Applied Linguistics 7*(2): 251-265.

Al-Zubeiry, H. 2015. ‘Linguistic analysis of Saudi pidginized Arabic as produced by Asian foreign expatriates.’ *International Journal of Applied Linguistics and English Literature 4*(2): 47-53.

Al-Masri, M. and A. Jongman. 2004. ‘Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results’. In *Proceedings of the 2003 Texas Linguistics Society Conference.* Somerville, MA: Cascadilla Proceedings Project. 96-106.

Amayreh, M. and A. Dyson. 1998. ‘The acquisition of Arabic consonants’. *Journal of Speech, Language, and Hearing Research 41*(3): 642-653.

Ambu Saidi, S. 2019. ‘A sociolinguistic study of dialect contact and change’. Unpublished Ph.D Dissertation, Newcastle University.

Anderson, J. 1987. ‘The markedness differential hypothesis and syllable structure difficulty’. In I. Georgette and S. Weinberger (eds) *Interlanguage Phonology: The Acquisition of A Second Language Sound System*. Cambridge, MA: Newbury House. 279-291.

Anderson, A. and B. Howarth. 2002. ‘September. Referential form and word duration in video-mediated and face-to-face dialogues’. In *Proceedings of the Sixth Workshop on the Semantics and Pragmatics of Dialogu.e* Edinburgh, UK. 4-6.

Andringa, S. and A. Gofroid. 2018. ‘SLA for all? Reproducing SLA research in non-academic samples’. Accessed December 13, 2018 from https://osf.io/mp47b/.

Andruski, J. and P. Kuhl. 1996. ‘The acoustic structure of vowels in mothers' speech to infants and adults’. Philadephia, USA. In *Proceeding of Fourth International Conference on Spoken Language Processing.* 1545-1548.

Aoyama, K., J. Flege, S. Guion, R. Akahane-Yamada and T. Yamada. 2004. ‘Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English/l/and/r’. *Journal of Phonetics 32*(2): 233-250.

Arthur, B., R. Weiner, M. Culver, Y. Lee and D. Thomas. 1980. ‘The register of impersonal discourse to foreigners: Verbal adjustments to foreign accent’. In D. Larsen-Freeman (ed.) *Discourse Analysis in Second Language Research*. Rowley: Newbury House. 111–124.

Asfoor, M. 1982. ‘Difficulties English speakers encounter in Arabic phonology’. Unpublished Ph.D Dissertation, University of San Francisco.

Avram, A. 2014. ‘Immigrant workers and language formation: Gulf Pidgin Arabic’. *Language and Migration 6(2)*: 7-40.

Baker, R. and V. Hazan. 2011. ‘DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs’. *Behavior Research Methods 43*(3): 761-770.

Bakir, M.J., 2010. ‘Notes on the verbal system of Gulf Pidgin Arabic’. *Journal of Pidgin and Creole Languages 25*(2): 201-228.

Baran, J., M. Zlatin Laufer and R. Daniloff. 1977. ‘Phonological contrastivity in conversation: A comparative study of voice onset time’. *Journal of Phonetics* 5: 339-350.

Barr, D., R. Levy, C. Scheepers and H. Tily. 2013. ‘Random effects structure for confirmatory hypothesis testing: Keep it maximal’. *Journal of Memory and Language 68*(3): 255-278.

Barriuso, T. and R. Hayes-Harb. 2018. ‘High Variability Phonetic Training as a Bridge from Research to Practice’. *CATESOL Journal 30*(1): 177-194.

Bassetti, B. 2007. ‘Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language’. In A. Guder, X Jiang and Y. Wan (eds) *The Cognition, Learning and Teaching of Chinese Characters*. Beijing: Beijing Language and Culture University Press. 155–179.

Bassetti, B., P. Escudero and R. Hayes-Harb. 2015. ‘Second language phonology at the interface between acoustic and orthographic input’. *Applied Psycholinguistics 36*(1): 1-6.

Bassetti, B., 2017. ‘Orthography affects second language speech: Double letters and geminate production in English’. *Journal of Experimental Psychology: Learning, Memory, and Cognition 43*(11): 1835.

Bates, E., 2003. ‘On the nature and nurture of language’. *Frontiers of Biology: The brain of Homo Sapiens*:241-265.

Bates, D., M. Maechler, B. Bolker and S. Walker. 2015. ‘Fitting linear mixed-effects models using l me4’. *Journal of Statistical Software* 67(1): 1-48.

Bateson, M. and K. Ryding. 1967. *Arabic Language Handbook*. Washington: Georgetown University Press.

Batibo, H. 2002. ‘The evolution of the Kiswahili syllable structure’. *South African Journal of African Languages 22*(1): 1-10.

Beckman, J., P. Helgason, B. McMurray and C. Ringen. 2011. ‘Rate effects on Swedish VOT: Evidence for phonological overspecification’. *Journal of Phonetics 39*(1): 39-49.

Beddor, P. and T. Gottfried. 1995. ‘Methodological issues in adult cross-language research’. In S. Winifred (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. York Press, Baltimore. 207-232.

Beebe, L. and H. Giles. 1984. ‘Speech-accommodation theories: A discussion in terms of second-language acquisition’. *International Journal of the Sociology of Language* 46: 5-32.

Belkaid, Y. 1984. ‘Arabic vowels, modern literature, spectrographic analysis’. *Phonetic Works Strasbourg Institution 16*: 217-240.

Bell, A., D. Jurafsky, E. Fosler-Lussier, C. Girand, M. Gregory, M. and D. Gildea. 2003. ‘Effects of disfluencies, predictability, and utterance position on word form variation in English conversation’. *The Journal of the Acoustical Society of America 113*(2): 1001-1024.

Belvins, J. 1995. ‘The syllable in phonological theory. In J. Goldsrnith (ed.) *The Handbook of Phonological Theory*. Cambridge, MS: Blackwell. 206-244.

Benders, T. 2013. ‘Mommy is only happy! Dutch mothers’ realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent’. *Infant Behavior and Development 36*(4): 847-862.

Behrens, S. and S. Blustein. 1988. ‘Acoustic characteristics of English voiceless fricatives: A descriptive analysis.’ *Phonetics 16*: 295-298.

Best, C., G. McRoberts, and N. Sithole. 1988. ‘Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants’. *Journal of Experimental Psychology: Human Perception and Performance 14*(3): 345.

Best, C. and W. Strange. 1992. ‘Effects of phonological and phonetic factors on cross-language perception of approximants’. *Journal of Phonetics 20*(3): 305-330.

Best, C. 1993. ‘Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development’. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage and J. Morton (eds) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht: Springer. 289-304.

Best, C. 1995. ‘A direct realist view of cross-language speech perception’. In W. Strange (ed.) *Speech Perception and Linguistic Experience*. Baltimore: York Press. 171-206.

Best, C. 1994. ‘The emergence of native-language phonological influences in infants: A perceptual assimilation model’. In J. Goodman and H Nusbaum (eds) *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Massachusets: The MIT Press.167-196.

Best, C., G. McRoberts and E. Goodell. 2001. ‘Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener’s native phonological system’. *The Journal of the Acoustical Society of America 109*(2): 775-794.

Best, C. and G. McRoberts. 2003. ‘Infant perception of non-native consonant contrasts that adults assimilate in different ways’. *Language and Speech 46*(2-3):183-216.

Best, C. and M. Tyler. 2007. ‘Nonnative and second-language speech perception: Commonalities and complementarities’. *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege 1334*: 1-47.

Bever, T. 1981. ‘Normal acquisition processes explain the critical period for language learning. In K. Diller (ed) *Individual Differences and Universals in Language Learning Aptitud*e. Rowley: Newbury House. 176– 198.

Bhaskararao, P. and A. Ray. 2017. ‘Telugu’. *Journal of the International Phonetic Association 47*(2): 231-241.

Biadsy, F., J. Hirschberg, and N. Habash. 2009. ‘Spoken Arabic dialect identification using phonotactic modeling’. In *Proceedings of the Workshop on Computational Approaches to Semitic Languages*. Association for Computational Linguistics. 53-61.

Bialystok, E. and K. Hakuta. 1999. ‘Confounded age: Linguistic and cognitive factors in age differences for second language acquisition’. In D. Birdsong (ed.) *Second Language Acquisition and the Critical Period Hypothesis*. London: Lawrence Erlbaum Associates Publishers. 161-181.

Biersack, S., V. Kempe, and L. Knapton. 2005. ‘Fine-tuning speech registers: a comparison of the prosodic features of child-directed and foreigner-directed speech’. In *Proceeding of Ninth European Conference on Speech Communication and Technology*. Lisbon, Portugal. 4-8.

Bigelow, M. and E. Tarone. 2004. ‘The role of literacy level in second language acquisition: Doesn't who we study determine what we know?’. *TESOL Quarterly 38*(4): 689-700.

Bin-Muqbil, M., 2006. ‘Phonetic and phonological aspects of Arabic emphatics and gutterals’. Unpublished Ph.D Dissertation, University of Wisconsin--Madison.

Birdsong, D. 1999. ‘Introduction: Whys and why nots of the critical period hypothesis for second language acquisition’. In D. Birdsong (ed.) *Second Language Acquisition and The Critical Period Hypothesis*. Mahwah: Lawrence Erlbaum Associates. 1-22.

Bizri, F. 2009. ‘Sinhala in contact with Arabic: The birth of a new pidgin in the Middle East’. *Annual Review of South Asian Languages and Linguistics*. 135-149.

Bizri, F. 2014. ‘Maids’ Talk’. *Colonization and Domestic Service: Historical and Contemporary Perspectives 14*: 113.

Bobb, S., K. Mello, E. Turco, L. Lemes, E. Fernandez and K. Rothermich. 2019. 'Second Language Learners' Listener Impressions of Foreigner-Directed Speech'. *Journal of Speech, Language, and Hearing Research 62*(9): 3135-3148.

Boersma, P. and D. Weenink. 2009. 'Praat: doing phonetics by computer'. Accessed September 30, 2016 from http://www.fon.hum.uva.nl/praat/.

Borden, G., K. Harris and L. Raphae. 1994. *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Baltimore: Lippincott Williams & Wilkins.

Bourhis, R. 1979. 'Language in ethnic interaction: A social psychological approach'. *Language and Ethnic Relations*: 117-141.

Bradlow, A., G. Toretta, D. Pisoni. 1996. 'Intelligibility of normal speech I: Global and fine- grained acoustic-phonetic talker characteristics'. *Speech Communication* 20: 255-272.

Bradlow, A. and T. Bent. 2002. 'The clear speech effect for non-native listeners'. *The Journal of the Acoustical Society of America 112*(1): 272-284.

Bradlow, A., N. Kraus and E. Hayes. 2003. 'Speaking clearly for children with learning disabilities'. *Journal of Speech, Language, and Hearing Research 46:* 80-97.

Bradlow, A. and J. Alexander. 2007. 'Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners'. *The Journal of the Acoustical Society of America 121*(4): 2339-2349.

Brame, M. 1974. 'Stress in Palestinian, Maltese and Spanish'. *Linguistic Inquiry 5*(1): 39-60.

Broesch, T. and G. Bryant. 2015. 'Prosody in infant-directed speech is similar across western and traditional cultures'. *Journal of Cognition and Development 16*(1): 31-43.

Brown, G. 1983. 'Prosodic structure and the given/new distinction.' In A. Cutler and D. Ladd (eds) *Prosody: Models and Measurements*. Berlin, Heidelberg: Springer. 67-77.

Broselow, E. 1983. 'Nonobvious transfer: On predicting epenthesis errors'. In S. Gass and G. Selinker (eds) *Language Transfer in Language Learning*. Rowley: Newbury House. 269-280.

Broselow, E. 1988. 'Prosodic phonology and the acquisition of a second language'. In S. Flynn and W. O'Neil (eds) *Linguistic Theory in Second Language Acquisition*. Dordrecht: Springer. 295-308.

Butcher, A. and K. Ahmad 1987. 'Some acoustic and aerodynamic characteristics of pharyngeal consonants in Iraqi Arabic'. *Phonetica 44*(3): 156-172.

Butler, S., L. O'Sullivan, B. Shah and N. Berthier. 2014. 'Preference for infant-directed speech in preterm infants'. *Infant Behavior and Development 37*(4): 505-511.

Bynon, J. 1968. 'Berber nursery language'. *Transactions of the Philological Society 67*(1): 107-161.

Cancino, H., E. Rosansky and J. Schumann. 1978. 'The acquisition of English negatives and interrogatives by native Spanish speakers'. *Second Language Acquisition: A book of Readings* 207-230.

Cairns, C. and M. Feinstein.1982. 'Markedness and the theory of syllable structure'. *Linguistic Inquiry*:193-225.

Card, E. 1983. A phonetic and phonological study of Arabic emphasis. Unpublished Ph.D Dissertation, Cornell University.

Carlisle, R. 1988. 'The Effect of Markedness on Epenthesis in Spanish/English Interlanguage Phonology'. *Issues and Developments in English and Applied Linguistics 3*: 15-23.

Carlisle, R.1991. 'The influence of environment on vowel epenthesis in Spanish/English interphonology'. *Applied Linguistics 12*(1): 76-95.

Carlisle, R. 2001. 'Syllable structure universals and second language acquisition'. *International Journal of English Studies 1*(1): 1-19.

Casagrande, J. 1948. 'Comanche baby language'. *International Journal of American Linguistics* 14: 11- 14.

Chafe, W. 1974. 'Language and consciousness'. *Language*: 111-133.

Chaudron, C. 1978. 'English as the Medium of Instruction in ESL Classes: An Initial Report of a Pilot Study of the Complexity of Teacher's Speech'. Toronto: Ontario Institute for Studies in Education.

Chaudron, C. 1988. *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.

Cho, T. and P. Ladefoged. 1999. 'Variation and universals in VOT: evidence from 18 languages'. *Journal of Phonetics 27*(2): 207-229.

Chomsky, N. 1981. 'Knowledge of language: Its elements and origins'. *Philosophical Transactions of the Royal Society of London. Biological Sciences 295*(1077): 223-234.

Christensen, R.H.B.2019. 'Regression Models for Ordinal Data R package ordinal version 2019.12-10].

Church, R., B. Bernhardt, K. Pichora-Fuller and R. Shi. 2005. 'Infant-directed speech: Final syllable lengthening and rate of speech'. *Canadian Acoustics 33*(4): 13-19.

Clahsen, H., J. Meisel and M. Pienemann M. 1983. *Deutsch als Zweitsprache: Der Spracherwerb ausländischer Arbeiter*. Tübingen: Gunter Narr Verlag.

Clahsen, H. 1991. *Child Language and Developmental Dysphasia: Linguistic Studies of the Acquisition of German*. Amsterdam: John Benjamins Publishing.

Clark, E. 2009. *First Language Acquisition*. Cambridge: Cambridge University Press.

Clements, G. 1990. 'The role of the sonority cycle in core syllabification'. In J. Kingston and M. Beckman (eds) *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press. 283-333.

Clyne, M. 1981. 'Second Generation' Foreigner Talk in Australia. *International Journal of the Sociology of Language 1981* (28): 69-80.

Cooper, R. and R. Aslin. 1990. 'Preference for infant-directed speech in the first month after birth.' *Child Development 61*(5): 1584-1595.

Cooper, R., J. Abraham, S. Berman, and M. Staska. 1997. 'The development of infants' preference for motherese'. *Infant Behaviour and Development* 20: 477–488.

Coupland, N., J. Coupland, H. Giles and K. Henwood. 1988. 'Accommodating the elderly: Invoking and extending a theory'. *Language in Society, 17*(1): 1-41.

Craats, I., J. Kurvers and M. Young-Scholten. 2006. 'Research on low-educated second language and literacy acquisition. *LOT Occasional Series 6*: 7-23.

Cristià, A. and A. Seidl. 2014. 'The hyperarticulation hypothesis of infant-directed speech'. *Journal of Child Language 41*(4): 913-934.

Cristià, A. 2010. 'Phonetic enhancement of sibilants in infant-directed speech'. *The Journal of the Acoustical Society of America 128*(1): 424-434.

Cross, J. 2002. 'Noticing'in SLA: Is it a valid concept. *TESL-EJ 6*(3): 1-9.

Cruttenden, A., 1994. 'Phonetic and prosodic aspects of Baby Talk'. In C. Gallaway and B. Ritchards (eds) *Input and Interaction in Language Acquisition.* Cambridge: Cambridge University Press. 135-152.

Cummins, J. 1991.' Interdependence of first-and second-language proficiency in bilingual children'. *Language Processing in Bilingual Children*: 70-89.

Cutler A. and D.Foss. 1977. 'On the role of sentence stress in sentence processing.' *Language and Speech 20*(1): 1-10.

Daland, R., B. Hayes, B. J. White, M. Garellek, A. Davis and I. Norrmann. 2011. 'Explaining sonority projection effects'. *Phonology 28*(2): 197-234.

Damer, N., M. Al-Ayyoub and I. Hmeidi. 2017. 'Automatically Determining Correct Application of Basic Quranic Recitation Rules'. In *Proceedings of the International Arab Conference on Information Technology*. Tunisia.

Dashti, A. 2013. 'Interacting with domestic workers in Kuwait: Grammatical features of foreigner talk. *International Journal of the Sociology of Language* (224): 63-84.

Davis, S., 1995. ‘Emphasis spread in Arabic and grounded phonology’. *Linguistic Inquiry 26*(3): 465-498.

Dejene G. and J. Devardhi. 2013. ‘Assimilation in Oromo Phonology’. *Language in India* 13(10): 331-358.

DeKeyser, R. and J. Larson-Hall. 2005. ‘What does the critical period really mean’. In J. Kroll and A. De Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press. 88-108.

Delattre, P. 1968. ‘From acoustic cues to distinctive features’. *Phonetica 18*(4): 198-230.

Delattre, P. 1971. ‘Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English’. *Phonetica 23*(3): 129-155.

Dela Rosa, J. and D. Arguelles. 2016. ‘Do Modification and interaction work? A critical review of literature on the role of foreigner talk in second language acquisition’. *Journal on English Language Teaching 6*(3): 46-60.

DeMatteo, D., G. Marczyk and D. Festinger. 2005. *Essentials of Research Design and Methodology*. New Jersey: John Wiley and Sons, Inc.

Dittmar, N. and W. Klein. 1979. ‘Developing grammars: The acquisition of German syntax by foreign workers: With 9 figures’. In W. Levelt (ed.) *Springer Series in Language and Communication*. Berlin: Springer.

Dodane, C. and J. Al-Tamimi. 2007. ‘An acoustic comparison of vowel systems in adult-directed-speech and child-directed speech: Evidence from French, English & Japanese’. In *16th International Congress of Phonetics Sciences*. 6-10.

Dragojevic, M., J. Gasiorek and H. Giles. 2016. ‘Accommodative strategies as core of the theory’. In H. Giles (ed.) *Communication Accommodation Theory: Negotiating Personal Relationships and Social Identities Across Contexts*. Cambridge: Cambridge University Press.36-59.

Dromey C. 2000. ‘Articulatory kinematics in patients with Parkinson disease using different speech treatment approaches’. *Journal of Medical Speech-Language Pathology* 8:155–161.

Dunst, C., E. Gorman and D. Hamby. 2012. ‘Preference for infant-directed speech in preverbal young children’. *Center for Early Literacy Learning 5*(1): 1-13.

Eady, S. and Cooper, W. 1986. ‘Speech intonation and focus location in matched statements and questions.’ *The Journal of the Acoustical Society of America 80*(2): 402-415.

Eaves Jr, B., N. Feldman, N., T. Griffiths, P. and Shafto. 2016. ‘Infant-directed speech is consistent with teaching’. *Psychological Review 123*(6): 758.

Eckman, F. ‘Markedness Differential Hypothesis (MDH)’. *Language 15*: 335-70.

Eckman, F. 1985. ‘Some Theoretical and Pedagogical Implications of the Markedness Differential Hypothesis’. *Studies in Second Language Acquisition* 7(3): 289-307.

Edwards, J. and M. Beckman. 2008. ‘Methodological questions in studying consonant acquisition’. *Clinical Linguistics and Phonetics 22*(12): 937-956.

Eimas, P., E. Siquel, P. Jusczyk and J. Vigorito. 1971.‘Speech perception in infants’. *Science* 171: 303–306.

Eimas, P. and J. Corbit. 1973. ‘Selective adaptation of linguistic feature detectors’. *Cognitive Psychology 4*(1): 99-109.

Eimas, P. 1975. ‘Speech perception in early infancy’. In L. Cohen and P. Salapatek (eds) *Infant Perception: From Sensation to Cognition*. New York: Academic Press. 193-231.

Ellis, R. 1985. *Understanding Second Language Acquisition.* Oxford: Oxford university press.

Ellis, R. 1991. ‘The interaction Hypothesis: A critical evaluation’. In *the Regional Language Center Seminar*. Singapore. 1-46.

Elgendy, A. 2001. *Aspects of Pharyngeal Coarticulation*. Amsterdam: LOT International Series.

Ellis, R.1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Englund, K. 2005. ‘Voice onset time in infant directed speech over the first six months’. *First language 25*(2): 219-234.

Englund, K. and D. Behne. 2006. ‘Changes in infant directed speech in the first six months’. *Infant and Child Development: An International Journal of Research and Practice 15*(2): 139-160.

Escudero, P. and P. Boersma. 2004. ‘Bridging the gap between L2 speech perception research and phonological theory’. *Studies in Second Language Acquisition 26*(4): 551-585.

Escudero, P. 2006. ‘The phonological and phonetic development of new vowel contrasts in Spanish learners of English’. *English with a Latin Beat: Studies in Portuguese/Spanish-English Interphonology 31*: 41.

Evans, M., 1987. ‘Linguistic accommodation in a bilingual family: One perspective on the language acquisition of a bilingual child being raised in a monolingual community’. *Journal of Multilingual & Multicultural Development 8*(3): 231-235.

Fant, G. 1960. *Acoustic Theory of Speech Perception*. The Hague: Mouton.

Fant, G. 1973. *Speech Sounds and Features*. Massachusetts: The MIT Press.

Fant, G., L. Nord and P. Branderud. 1976. ‘A note on the vocal tract wall impedance’. *Speech Transmission Laboratory Quarterly Progress and Status Report 4*: 13-27.

Ferguson, C. 1956. ‘Arabic baby talk’. In M. Halle (ed.) *For Roman Jakobson: Essays in the Occasion of his 60th Birthday*. Mouton: The Hague. 121-128.

Ferguson, C. and M. Chowdhury. 1960. ‘The phonemes of Bengali’. *Language 36*(1): 22-59.

Ferguson, C. 1964. ‘Baby talk in six languages’. *American Anthropologist 66*(6): 103-114.

Ferguson, C. 1971. ‘Absence of copula and the notion of simplicity’. *Pidginization and Creolization of Languages*: 141-150.

Ferguson, C. 1975. ‘Toward a characterization of English foreigner talk’. *Anthropological Linguistics*: 1-14.

Ferguson, C. 1977. ‘Baby talk as a simplified register’. In C. Snow and C. Ferguson (eds) *Talking to Children*. Cambridge: Cambridge University Press. 209-235.

Ferguson, C. 1981. ‘Foreigner talk’as the name of a simplified register. *International Journal of the Sociology of Language 1981*(28): 9-18.

Ferguson, S. and D. Kewley-Port. 2002. ‘Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners’. *The Journal of the Acoustical Society of America 112*(1): 259-271.

Ferguson, S. and D. Kewley-Port. 2002. ‘Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners’. *Journal of Acoustic Society of America* 112: 259-271.

Fernald, A. and T. Simon. 1984. ‘Expanded intonation contours in mothers' speech to newborns’. *Developmental Psychology 20*(1): 104-113.

Fernald, A. and P. Kuhl. 1987. ‘Acoustic determinants of infant preference for motherese speech’. *Infant behavior and development 10*(3): 279-293.

Fernald, A., T. Taeschner, J. Dunn, M. Papousek, B. de Boysson-Bardies and I. Fukui. 1989. ‘A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants’. *Journal of Child Language 16*(3): 477-501.

Fernald, A. and C. Mazzie. 1991. ‘Prosody and focus in speech to infants and adults.’ *Developmental Psychology 27*(2): 209-221.

Fischer, W., 2013. ‘Classical Arabic’. In R. Hetzron (ed.) *The Semitic Languages. London: Routledge.* 187-219.

Flege, J. 1981. ‘The phonological basis of foreign accent: A hypothesis’. *TESOL Quarterly 15*(4): 443-455.

Flege, J. and R. Port. 1981. ‘Cross-language phonetic interference: Arabic to English’. *Language and Speech* 24 (2): 125-146.

Flege, J. and R. Davidian. 1984. ‘Transfer and developmental processes in adult foreign language speech production’. *Applied Psycholinguistics 5*(4): 323-347.

Flege, J. 1987. ‘Effects of equivalence classification on the production of foreign language speech sounds’. In A. James and A. Leather (eds) *Sound Patterns in Second Language Acquisition*. Dordrecht: Foris. 9-39.

Flege, J. 1988. ‘Factors affecting degree of perceived foreign accent in English sentences’. *The Journal of the Acoustical Society of America 84*(1): 70-79.

Flege, J. 1993. ‘Production and perception of a novel, second-language phonetic contrast’. *The Journal of the Acoustical Society of America 93*(3): 1589-1608.

Flege, J. 1995. ‘Second language speech learning: Theory, findings, and problems’. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research 92*: 233-277.

Flege, J., M. Munro and I. MacKay.1995. ‘Factors affecting strength of perceived foreign accent in a second language’. *The Journal of the Acoustical Society of America 97*(5): 3125-3134.

Flege, J., O. Bohn and S. Jang.1997. ‘Effects of Experience on Non-native Speakers' Production and Perception of English Vowels’. *Journal of Phonetics* 25(4): 437-70.

Flege, J. 1999. ‘The relation between L2 production and perception’. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granveille and A. Bailey (eds) *Proceedings of the XIVth International Congress of Phonetics Sciences*. San Francisco. 1273-1276.

Flege, J. 1999. ‘Age of learning and second language speech’. In D. Birdsong (ed.) *Second Language Acquisition and The Critical Period Hypothesis*. Mahwah: Lawrence Erlbaum Associates Publishers. pp. 111-142.

Flege, J., G. Yeni-Komshian and S. Liu. 1999. ‘Age constraints on second-language acquisition’. *Journal of Memory and Language 41*(1): 78-104.

Flege, J. and S. Liu. 2001. ‘The effect of experiemce on adults’ acquisition of a second language’. *Studies in Second Language Acquisition 23*(4): 527-552.

Flege, J., 2002. ‘*Interactions between the native and second-language phonetic systems’.* In P. Burmeister, T. Piske and A. Rohde (eds) *An Integrated View of Language Development: Papers in Honor of Henning Wode*. 217-244.

Flege, J. 2003. ‘Assessing constraints on second-language segmental production and perception’. *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities 6*: 319-355.

Flege, J. 2009. ‘Give input a chance’. In T. Piske and M. Young-Scholten (eds) *Input matters in SLA*. Bristol: Multilingual Matters. 175-190.

Fowler, C. and J. Housum. 1987. ‘Talkers' signaling of “new” and “old” words in speech and listeners' perception and use of the distinction’. *Journal of Memory and Language 26*(5): 489-504.

Fowler, C. 1988. Differential shortening of repeated content words produced in various communicative contexts. *Language and speech 31*(4): pp.307-319.

Forrest, K., G. Weismer, P. Milenkovic, and R. Dougall. 1988. ‘Statistical Analysis of Word-initial Voiceless Obstruents: Preliminary Data’. *The Journal of the Acoustical Society of America* 84.1: 115-23.

Freed, B. 1981. ‘Foreigner talk, baby talk, native talk’. *International Journal of the Sociology of Language* 28: 19-40.

French, K. 1988. ‘Insights into Tagalog: Reduplication, Infixation, and Stress from Nonlinear Phonology’. The Summer Institute of Linguistics and the University of Texas at Arlington.

Fries, C. 1957. Preface to R. Lado’s *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.

Fuchs, S. and P. Birkholz P. 2019. ‘Phonetics of Consonants’. In *Oxford Research Encyclopedia of Linguistics*.

Furrow, D., K. Nelson and H. Benedict. 1979. ‘Mothers' speech to children and syntactic development: Some simple relationships’. *Journal of Child Language 6*(3): 423-442.

Gahl, S., Y. Yao and K. Johnson. 2012. ‘Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech’. *Journal of Memory and Language 66*(4): 789-806.

Gallois, C., T. Ogay and H. Giles. 2005. ‘Communication accommodation theory’. In W. Gundykunst (ed.) *Theorizing about Intercultural Communication*. Thousand Oaks: Sage. 121-148.

Gass, S. and E. Varonis. 1984. ‘The effect of familiarity on the comprehensibility of non-native speech’. *Language Learning* 34(11): 65-89.

Gass, S. and E. Varonis. 1985. ‘Variation in native speaker speech modification to non-native speakers’. *Studies in Second Language Acquisition 7*(1): 37-57.

Gass, S. 1988. ‘Integrating research areas: A framework for second language studies1’. *Applied Linguistics 9*(2): 198-217.

Gass, S., 2003. ‘Input and interaction’. In C. Doughty and M. Long (eds) *The Handbook of Second Language Acquisition 2*. Georgetown: Georgetown University Press. 224-256

Gass, S. and A. Mackey. 2007. 'Input, interaction, and output in second language acquisition'. In B. Van Patten and J. Williams (eds) *Theories in Second Language Acquisition: An Introduction*. New York: Routledge. 180-206.

Gass, S. 2009. 'Second language acquisition.' In S. Foster-Cohen (ed.) *Language Acquisition*. Palgrave Macmillan: London. 109-139.

Gerrits, E. and M. Schouten. 2004. 'Categorical perception depends on the discrimination task'. *Perception and Psychophysics 66*(3): 363-376.

Chaudron, C. 1988. *Second Language Classrooms: Research on Teaching and Learning*. Cambridge: Cambridge University Press.

Ghazeli, S. 1977. 'Back Consonants and Backing Coarticulation in Arabic'. Unpublished Doctoral Dissertation, University of Texas at Austin.

Ghorbanian, M. and S. Jabbarpoor. 2017. 'Effects of lexical simplification and elaboration on Iranian intermediate EFL learners' learning and retention of phrasal verbs'. *Theory and Practice in Language Studies 7*(11): 1003-1010.

Ghosal, S. and S. Porkodi, 2014. 'Reasons for migration of contract labour to Oman: A SEM approach'. *International Journal of Management and Business Research 4*(3): 213-224.

Giannini, A. and M. Pettorino. 1982. 'The emphatic consonants in Arabic'. *Speech Laboratory Report IV*. Napoli: Istituto Universitario Orientale di Napoli.

Giles, H., N. Coupland, and J. Coupland. 1991. 'Accommodation theory: Communication, context, and consequence'. In *Contexts of Accommodation: Developments in Applied Sociolinguistics*. *1-68*.

Giles, H. and T. Ogay. 2007. Communication accommodation theory. In B. Whaley and W. Samter (ed) *Explaining Communication: Contemporary Theories and Exemplars*. Mahwah, NJ: Lawrence Elbraum. 293-310.

Giles and Gasiorek. 2013. 'Parameters of Nonaccommodation: Refi ning and Elaborating Communication Accommodation Theory'. In J. Forgas, O. Vincze and J. László (eds) *Social Cognition and Communication*. Psychology Press. 169-186

Gleitman, L., E. Newport and H. Gleitman, H. 1984. 'The current status of the motherese hypothesis'. *Journal of Child Language 11*(1): 43-79.

Gleitman, L. and E. Wanner. 1982. *Language Acquisition: The State of the Art*. Cambridge: Cambridge University Press.

Godfrey, M., M. Ruhs, N. Shah and M. Smith. 2004. 'Migrant domestic workers in Kuwait: Findings based on a field survey and additional research'. In S. Esim and M. Smith (eds) *Gender and Migration in Arab States: The Case of Domestic Workers*. Regional Office for Arab States. Beirut. 42-63.

Gordon, M., P. Barthmaier and K. Sands. 2002. 'A cross-linguistic acoustic study of voiceless fricatives'. *Journal of the International Phonetic Association 32*(2): 141-174.

Greenberg, J. 1978. 'Some generalizations concerning initial and final consonant clusters'. *Universals of human language 2*: 243-279.

Grieser, D. and P. Kuhl. 1988. 'Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese'. *Developmental Psychology 24*(1):14-20.

Grosjean, F. and J. Gee.1987. 'Prosodic structure and spoken word recognition'. *Cognition 25*(1-2): 135-155.

Guenther, F.1994. 'A neural network model of speech acquisition and motor equivalent speech production'. *Biological Cybernetics 72*(1): 43-53.

Guion, S., J. Flege, S. Liu, and G. Yeni-Komshian. 2000. 'Age of learning effects on the duration of sentences produced in a second language'. *Applied Psycholinguistics 21*(02): 205-228.

Haddad, G. F. 1984. 'Epenthesis and sonority in Lebanese Arabic'. *Studies in the Linguistic Sciences* 14(1): 57-88.

Halle, M., G. Hughes and J.Radley. 1957. 'Acoustic properties of stop consonants'. *The Journal of the Acoustical Society of America 29*(1): 107-116.

Halliday, M. 1967. 'Notes on transitivity and theme in English Part I'. *Journal of Linguistics 3*(1): 37-81.

Hamid, A., Al-Saqqaf, A. and M. Al-Mashani. 2008. 'The Nizwa dialect: A phonological study'. *Magazine of Linguistic Studies 10*(2): 28-116.

Hansen, J.G., 2004. 'Developmental sequences in the acquisition of English L2 syllable codas: A preliminary study'. *Studies in Second Language Acquisition 26*(1): 85-124.

Harder, P. 1980. 'Discourse as self-expression—on the reduced personality of the second-language learner'. *Applied Linguistics 1*(3): 262-270.

Hartman, K., N. Ratner and R. Newman. 2017. 'Infant-directed speech (IDS) vowel clarity and child language outcomes'. *Journal of Child Language 44*(5): 1140-1162.

Hatch, E., R. Sapira and J. Gough. 1975. 'Foreigner talk discousre'. *Working Papers in English as a Second Language*. Los Angeles: U.C.L.A.

Hatch, E. 1983. ‘Simplified input and second language acquisition’. *Pidginization and creolization as Language Acquisition*: 64-86.

Hassan, Z. and J. Esling. 2007. ‘Laryngoscopic (articulatory) and acoustic evidence of a prevailing emphatic feature over the word in Arabic’. In *International Congress of Phonetic Sciences* 16: 1753-1756.

Hayes-Harb, R. and K. Durham. 2016. ‘Native English Speakers’ Perception of Arabic Emphatic Consonants and the Influence of Vowel Context’. *Foreign Language Annals 49*(3): 557-572.

Hazan, V., M. Uther and S. Granlund. 2015. ‘How does foreigner-directed speech differ from other forms of listener-directed clear speaking styles?’. In *Proceedings of the 18th International Congress of Phonetic Sciences*.

Henrich, J., S. Heine and A. Norenzayan. 2010. ‘Most people are not WEIRD’. *Nature 466*(7302): 29-29.

Henzl, V. 1979. ‘Foreign talk in the classroom’. *International Review of Applied Linguistics 17*(2): 159-167.

Heselwood, B., 1996. ‘Glottal states and emphasis in Baghdadi and Cairene Arabic: synchronic and diachronic aspects’. *University of Durham Center for Middle Eastern and Islamic Studies, Occasional Paper 53*.

Heselwood, B. 2007. ‘The ‘tight approximant’ variant of the Arabic ‘ayn’. *Journal of the International Phonetic Association 37*(1): 1-32.

Hillenbrand, J., M. Clark and R. Houde. 2000. ‘Some effects of duration on vowel recognition’. *The Journal of the Acoustical Society of America 108*(6): 3013-3022.

Hillerich, R. 1976. ‘Toward an assessable definition of literacy’. *The English Journal 65*(2): 50-55.

Hoberman, R. 1995. ‘Current Issues in Semitic Phonology’. In J. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cambridge: Blackwell. 839-47.

Holes, C. 2004. *Modern Arabic: Structures, functions, and varieties*. Washington DC: Georgetown University Press.

Huang, B. 2014. ‘The effects of age on second language grammar and speech production’. *Journal of Psycholinguistic Research 43*(4): 397-420.

Hughes, G. and M. Halle. 1956. ‘Spectral properties of fricative consonants.’ *The Journal of the Acoustical Society of America 28*(2): 303-310.

Hyman, L. and F. Katamba. 1999. ‘The syllable in Luganda phonology and morphology’. In H. Hulst and N. Ritter (eds) *The Syllable: Views and Facts. Berlin: Mouton de Gruyter*. 349-416.

Ingram, D. 1989. *First Language Acquisition: Method, description and Explanation*. Cambridge: Cambridge University Press.

Ioup, G., E. Boustagui, M. El Tigi and M. Moselle. 1994. ‘Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment’. *Studies in second language acquisition 16*(1): 73-98.

Israel, A., M. Proctor, L. Goldstein, K. Iskarous and S. Narayanan. 2012. ‘Emphatic segments and emphasis spread in Lebanese Arabic: a real-time magnetic resonance imaging study’. In *Thirteenth Annual Conference of the International Speech Communication Association*.

Iverson, P. and P. Kuhl, 1996. ‘Influences of phonetic identification and category goodness on American listeners’ perception of/r/and/l/’. *The Journal of the Acoustical Society of America 99*(2): 1130-1140.

James, A.1988. ‘Patterns in the phonology of second language acquisition’. *In The acquisition of a second language phonlogy: A linguistic theory of developing sound structures.* Dischingerweg: Gunter Narr VerlagTübingen. 2-75.

Janicki, K. 1986. ‘Accommodation in native speaker-foreigner interaction’. In J. House and S. Blum-Kulka (eds) *Interlingual and Intercultural Communication: Discourse and Cognition in Translation and Second Language Acquisition Studies.* Tübingen: Gunter Narr Verlag.

Janko, E., E. Dąbrowska and J. Street. 2019. Education and input as predictors of second language attainment in naturalistic contexts. *Languages 4*(3): 70.

Jing, C., Y. Chunsheng and L. Guofa. 2019. ‘Analysis of Second Language Acquisition (SLA) Speech Perception Model & the Perception of Second Language Prosody’. *Revista de Cercetare si Interventie Sociala*, *64*: 334-351.

Johnson, K. 2005. ‘Speaker normalization in speech perception’. In D. Pisoni and R. Remez (eds) *The handbook of Speech Perception.* Oxford: Blackwell Publishers. 363–89.

Jones, D. 1960. An Outline of English Phonetics. Cambridge: Heffer and Sons.

Jongman, A., R. Wayland and S. Wong. 2000. ‘Acoustic characteristics of English fricatives’. *The Journal of the Acoustical Society of America 108*(3): 1252-1263.

Jongman, A., W. Herd and M. Al-Masri. 2007. ‘Acoustic correlates of emphasis in Arabic’. In *Proceedings of the 16th International Congress of Phonetic Science.* 913-316.

Jongman, A., W. Herd, M. Al-Masri, J.Sereno and S. Combest. 2011. ‘Acoustics and perception of emphasis in Urban Jordanian Arabic’. *Journal of Phonetics 39*(1): 85-95.

Joos, M. 1948. ‘Acoustic phonetics’. *Language 24*(2): 5-136.

Jurafsky, D., A. Bell, E. Fosler-Lussier, C. Girand and W. Raymond. 1998. ‘Reduction of English function words in Switchboard’. In *Proceeding of Fifth International Conference on Spoken Language Processing 98*. Sydney. 1-4.

Jusczyk, P., A. Cutler and N. Redanz. 1993. ‘Infants' preference for the predominant stress patterns of English words’. *Child Development 64*(3): 675-687.

Kambuziya, A. 2007. ‘Arabic definite prefix: An autosegmental analysis’. *South Asian Language Review 17*(2): 63-73.

Katamba, F. 1974. ‘Aspects of the grammar of Luganda’. Unpublished Doctoral Dissertation, University of Edinburgh.

Katz, J. 1977. Foreigner Talk Input in Child Second Language Acquisition: Its Form and Function over Time.

Katz, J. 1981. ‘Children’s second-language acquisition: The role of foreigner talk in child-child interaction’. *International Journal of the Sociology of Language 1981*(28): 53-68.

Kang, K. and S. Guion. 2008. ‘Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies’. *The Journal of the Acoustical Society of America 124*(6): 3909-3917.

Kelkar, A. 1964. ‘Marathi baby talk’. *Word 20*(1): 40-54.

Kenstowicz, M. 1994. *Phonology in Generative Grammar*. Cambridge: Blackwell.

Kent, R.and C. Read. 2002. *The Acoustic Analysis of Speech.* Canada: Singular Thomson Learning. 139-88.

Khattab, G. 2002. ‘VOT production in English and Arabic bilingual and monolingual children’. *In D.* Parkinson and E. Benmamoun (ed.) *Perspectives on Arabic Linguistics XIII-XIV: Papers from the Thirteenth and Fourteenth Annual Symposia on Arabic Linguistics*. Amsterdam/ Philadelphia: John Benjamins Publishing. 1-38.

Khattab, G. and J. Al-Tamimi. 2008. ‘Durational cues for gemination in Lebanese Arabic’. *Language and Linguistics 22*: 39-55.

Kiparsky, P. 1993. ‘An OT perspective on phonological variation’. *Handout from Rutgers Optimality Workshop*. Accessed April 13, 2018 from https://web.stanford.edu/~kiparsky/#handouts.

Kirchhoff, K. and S. Schimmel. 2005. ‘Statistical properties of infant-directed versus adult-directed speech: Insights from speech recognition’. *The Journal of the Acoustical Society of America 117*(4): 2238-2246.

Kleifgen, J. 1985. ‘Skilled variation in a kindergarten teacher's use of foreigner talk’. *Input in Second Language Acquisition*: 59-68.

Klein, W. and C. Perdue. 1997. 'The Basic Variety (or: Couldn't natural languages be much simpler?)'. *Second Language Research 13*(4): 301-347.

Knoll, M. and M. Uther. 2004. 'Motherese and Chinese: Evidence of acoustic changes in speech directed at infants and foreigners'. *The Journal of the Acoustical Society of America 116*(4): 2522-2522.

Knoll, M. and L. Scharrer. 2007. 'Acoustic and affective comparisons of natural and imaginary infant-foreigner and adult-directed speech'. In *Eighth Annual Conference of the International Speech Communication Association*.1414-1417.

Knoll, M. and A. Costall. 2015. 'Characterizing F(0) contour shape in infant-and foreigner-directed speech'. *Speech Communication 66*: 231-243.

Knoll, M., M. Johnstone and C. Blakely. 2015. 'Can you hear me? Acoustic modifications in speech directed to foreigners and hearing-impaired people'. In *Sixteenth Annual Conference of the International Speech Communication Association*. Germany. 2987-2990.

Ko, E. 2012. 'Nonlinear development of speaking rate in child-directed speech'. *Lingua 122*(8): 841-857.

Krashen, S. 1982. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press Inc.

Krause, J. and L. Braida. 1995. 'The effects of speaking rate on the intelligibility of speech for various speaking modes'. *The Journal of the Acoustical Society of America 98*(5): 2982-2982.

Kuhl, P. 1991. 'Human adults and human infants show a 'perceptual magnetic effect' for the prototypes of speech categories, monkeys do not'. *Perception & Psychophysics 50*: 93-107.

Kuhl, P. and P. Iverson, P. 1995. 'Linguistic Experience and the perceptual magnet effect'. In: W. Strange (ed.) *Speech Perception and Linguistic Experience*. Timonium, MD: York Press. 121-154.

Kuhl, P. 2000. 'A new view of language acquisition'. In *Proceedings of the National Academy of Sciences 97*(22): 11850-11857.

Kuhl, P., K. Williams, F. Lacerda, K. Stevens and B. Lindblom. 1992. 'Linguistic experience alters phonetic perception in infants by 6 months of age'. *Science 255*(5044): 606-608.

Kuhl, P., J. Andruski, I. Chistovich, L. Chistovich, E. Kozhevnikova, V. Ryskina, E. Stolyarova, U. Sundberg, and F. Lacerda. 1997. 'Cross-language analysis of phonetic units in language addressed to infants'. *Science 277*(5326): 684-686.

Kuhl, P., B. Conboy, D. Padden, T. Nelson and J. Pruitt. 2005. ‘Early speech perception and later language development: Implications for the" critical period"’. *Language Learning and Development 1*(3-4): 237-264.

Kuhl, P., E. Stevens, A. Hayashi, T. Deguchi, S. Kiritani and P. Iverson. 2006. ‘Infants show a facilitation effect for native language phonetic perception between 6 and 12 months’. *Developmental Science 9*(2): F13-F21.

Kuhl, P. 2007. ‘Cracking the speech code: How infants learn language’. *Acoustical Science and Technology 28*(2): 71-83.

Kurvers, J., T. Vallen, and R. Hout. 2006. ‘Discovering features of language: Metalinguisitc awareness of adult illiterates’. *LOT Occasional Series 6*: 69-88.

Kuriyagawa, F., 1984. ‘The features of/ k/ and/ q/ in Cairo Standard Arabic’. *Annual Bulletin 18.* Research Institute of Logopedics and Phoniatrics, University of Tokyo. 65-73.

Kuznetsova A, P. Brockhoff and R. Christensen. 2017. ‘lmerTest package: Tests in linear mixed eff ects Models’. *Journal of Statistical Software* 82(13): 1-26.

Lababidi, Z., 2016. ‘The L2 perceptual mapping of Arabic and English consonants by American English learners’. Unpublished Ph.D Dissertation, The University of Wisconsin-Milwaukee.

Ladefoged, P., R. Glick and C. Criper. 1972. *Language in Uganda*. Oxford: Oxford University Press.

Ladefoged, P. and K. Johnson. 2001. *A Course in Phonetics*. Australia: Cengage Learning.

Lado, R. 1957. *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor, MI: The University of Michigan Press.

Lam, J., K. Tjaden, and G. Wilding. 2012. ‘Acoustics of clear speech: Effect of instruction’. *Journal of Speech, Language, and Hearing Research 55(6)*: 1807-1821.

Laradi, W.1983. ‘Pharyngealization in Libyan (Tripoli) Arabic: An instrumental study’. Unpublished Doctoral Dissertation, University of Edinburgh.

Laufer, A. and I. Condax. 1979. ‘The epiglottis as an articulator’. *Journal of the International Phonetic Association 9*(2): 50-56.

Laufer, A. and T. Baer. 1988. ‘The emphatic and pharyngeal sounds in Hebrew and in Arabic’. *Language and Speech 31*(2): 181-205.

Laufer, A. 1996. ‘The common [ʕ] is an approximant and not a fricative’. *Journal of the International Phonetic Association 26*(2): 113-118.

Laver, J. 1980. ‘The phonetic description of voice quality’. *Cambridge Studies in Linguistics London, 31*: 1-186.

Leather, J. and A. James. 1996. 'The acquisition of second language speech'. In W. Ritchie and T. Bhatia (eds) *Handbook of Second Language Acquisition*. San Diego: Academic Press. 269-316.

Lee, K.1992. 'Context-Dependent Phonetic Hidden Markov Models for Speaker-Independent Continuous Speech Recognition'. In P. Laface and R. De Mori (eds) *Speech Recognition and Understanding*. Berlin: Springer. 133-133.

Lehiste, I. and G. Peterson. 1961. 'Transitions, glides, and diphthongs'. *The journal of The Acoustical Society of America 33*(3): 268-277.

Lehiste, I. 1976. 'Suprasegmental features of speech'. In N. Lass (ed.) *Contemporary Issues in Experimental Phonetics*. New York: Academic Press. *225*-239.

Lenneberg, E. 1967. 'The biological foundations of language'. *Hospital Practice 2*(12): 59-67.

Lenth, R. and M. Lenth. 2018. 'Package 'lsmeans''. *The American Statistician 34*(4): 216-221.

Levelt, C. and R. Van de Vijver. 2004. 'Syllable types in cross-linguistic and developmental grammars'. In R. Kager, J. Pater, and W. Zonneveld (eds) *Constraints in Phonological Acquisition*. New York: Cambridge University Press. 204-218.

Levin, H., C. Snow and K. Lee. 1984. 'Nurturan talk to children. *Language and Speech* 27(2): 147-162.

Lieberman, P. and S. Blumstein. 1988. *Speech Physiology, Speech Perception, and Acoustic Phonetics*. Cambridge: Cambridge University Press.

Liénard, J. and M. Di Benedetto.1999. 'Effect of vocal effort on spectral properties of vowels'. *The Journal of the Acoustical Society of America 106*(1): 411-422.

Lindblom, B. 1990. 'Explaining phonetic variation: A sketch of the H&H theory'. In W. Hardcastle and A. Marchal (eds) *Speech Production and Speech Modelling.* Netherlands: Springer. 403-439.

Lindblom, B., 1996. 'Role of articulation in speech perception: Clues from production'. *The Journal of the Acoustical Society of America 99*(3): 1683-1692.

Lisker, L. and A. Abramson. 1964. 'A cross-language study of voicing in initial stops: Acoustical measurements'. *Word 20*(3): 384-422.

Liu, H, P. Kuhl, F. Tsao. 2003. 'An association between mothers' speech clarity and infants' speech discrimination skills'. *Developmental Science* 6(3): F1- F10.

Llamzon, T., 1966. 'Tagalog phonology'. *Anthropological Linguistics*: 30-39.

Lloret, M. 1995. 'The representation of glottals in Oromo". *Phonology 12*(2): 257-280.

Lloret, M. 1997.' Oromo phonology'. In A. Kaye and P. Daniels (eds) *Phonologies of Asia and Africa* 1. Indiana: Eisenbrauns. 493-520.

Long, M. 1981. ‘Input, interaction, and second-language acquisition’. *Annals of the New York Academy of Sciences 379*(1): 259-278.

Long, M., 1983a. ‘Linguistic and conversational adjustments to non-native speakers’. *Studies in Second Language Acquisition 5*(2): 177-193.

Long, M., 1983b. ‘Native speaker/non-native speaker conversation and the negotiation of comprehensible input’. *Applied Linguistics 4*(2): 126-141.

Long, M. 1985. Input and second language acquisition theory. In S. Gass and C. Madden (eds) *Input in Second Language Acquisition*. Rowley, Mass: Newbury House. 377-393.

Long, M. and P. Porter, 1985. ‘Group work, interlanguage talk, and second language acquisition.’ *TESOL Quarterly 19*(2): 207-228.

Long, M. 1990. ‘Maturational constraints on language development’. *Studies in Second Language Acquisition 12*(3): 251-285.

Lombardi, L. 2003. ‘Second language data and constraints on manner: Explaining substitutions for the English interdentals’. *Second Language Research 19*(3): 225-250.

Lynch, T. 2001. ‘Seeing what they meant: Transcribing as a route to noticing’. *ELT Journal 55*(2): 124-132.

Lynch, A. 2009. ‘Sociolinguistic Analysis of final /s/ in Miami Cuban Spanish’. *Language Sciences 31*: 766-90.

Macmillan, N. and C. Creelman. 1991. *Detection Theory: A User’s Guide*. Cambridge: Cambridge University Press.

Maddieson, I. and S. Disner. 1984. *Patterns of Sounds*. Cambridge: Cambridge University press.

Major, R. and M. Faudree. 1996. ‘Markedness universals and the acquisition of voicing contrasts by Korean speakers of English’. *Studies in Second Language Acquisition 18*(1): 69-90.

Maleki, Z. and A. Pazhakh. 2012. ‘The Effects of Pre Modified Input, Interactionally Modified Input, and Modified Output on EFL Learners' Comprehension of New Vocabularies’. *International Journal of Higher Education 1*(1): 128-137.

Malsheen, B. 1980. ‘Two hypotheses for phonetic clarification in the speech of mothers to children’. *Child Phonology 2*: 173-184.

Maniwa, K., A. Jongman and T. Wade. 2009. ‘Acoustic characteristics of clearly spoken English fricatives’. *The Journal of the Acoustical Society of America 125*(6): 3962-3973.

Mason, C. and W. Perreault Jr. 1991. ‘Collinearity, power, and interpretation of multiple regression analysis’. *Journal of Marketing Research 28*(3): 268-280.

Meisel, J. 1977. ‘Linguistic simplification: A study of immigrant workers' speech and foreigner talk’. In S. Corder and E. Roulet (eds) *The Notion of Simplification, Interlanguages and Pidgins and Their Relation to Second Language Pedagogy*. Geneve: Droz. 88-113.

Messer, D. 1981. ‘The identification of names in maternal speech to infants’. *Journal of Psycholinguistic Research 10*(1): 69-77.

McAllister, R. 2001. Experience as a factor in L2 phonological acquisition. *Working Papers-Lund University Department of Linguistics*. 116-119.

McCarthy, J. 1994. ‘The phonetics and phonology of Semitic pharyngeals’. In P. Keating (ed.) *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form* 86. Cambridge: Cambridge University Press. 191-233.

McCloy. D. 2016. ‘PhonR: tools for phoneticians and phonologists’. *R Package Version* 1.0-7.

McMurray, B., K. Kovack-Lesh, D. Goodwin and W. McEchron. 2013. ‘Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence?’. *Cognition 129*(2): 362-378.

Midi, H., S. Sarkar and Rana. 2010. ‘Collinearity diagnostics of binary logistic regression model’. *Journal of Interdisciplinary Mathematics 13*(3): 253-267.

Miller, C. 2004. ‘Impact of migration on Arabic urban vernacular: Advocating a comparative analysis. In *Proceedings of the 5th International Aida Conference*. Cadiz: Spain. 251-262.

Milroy, L. 1980. ‘Social network and language maintenance’. In B. Mayor and A. Pugh (eds) *Language, Communication and Education*. New Hapshire: Library Congress Cataloging. 70-81.

Mitchell, R. 2001. ‘Americans’ talk with dogs: Similarities and differences with talk to infants’. *Research on Language and Social Interaction* 34(2): 183-210.

Mitchell, R. 2004. ‘Controlling the dog, pretending to have a conversation, or just being friendly?: Influences of sex and familiarity on Americans’ talk to dogs during play’. *Interaction Studies 5*(1): 99-129.

Miyazawa, K., T. Shinya, A. Martin, H. Kikuchi and R. Mazuka. 2017. ‘Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition 166*: 84-93.

Mohamed, M. 2001. *Modern Swahili Grammar*. Nairobi: East Educational African Publishers.

Monsen, R., A. Engebretson and N. Vemula. 1978. ‘Indirect Assessment of the Contribution of Subglottal Air Pressure and Vocal-fold Tension to Changes of Fundamental Frequency in English’. *The Journal of the Acoustical Society of America* 64(1): 65-80.

Morais, J., L. Cary, J. Alegría, P. Bertelson. 1979. ‘Does awareness of speech as a sequence of phones arise spontaneously?’. *Cognition* 7: 323–331.

Morais, J., P. Bertelson, L. Cary and J. Alegria. 1986. ‘Literacy training and speech segmentation’. *Cognition 24*(1-2): 45-64.

Moyer, A. 2004. *Age, Accent, and Experience in Second Language Acquisition: An Integrated Approach to Critical Period Inquiry*. Bristol: Multilingual Matters.

Moyer, A. 2009. ‘Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment’. In P. Thorsten and M. Young-Scholten *Input Matters in SLA*. Buffalo: Multilingual Matters. 159-174.

Moyer, A. 2013. *Foreign Accent: The Phenomenon of Non-Native Speech*. Cambridge: Cambridge University Press.

Muñoz, C. 2008. ‘Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning’. *Applied Linguistics 29*(4): 578-596.

Muhammad, A. ul Qayyum, S. Tanveer, A. Martinez-Enriquez and A. Syed. 2012. ‘E-hafiz: Intelligent system to help Muslims in recitation and memorization of Quran’. *Life Science Journal* 9(1): 534-541.

Munro, M. 1993. ‘Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings’. *Language and Speech 36*(1): 39-66.

Munro, M. and T. Derwing. 1995. ‘Foreign accent, comprehensibility, and intelligibility in the speech of second language learners’. *Language Learning 45*(1): 73-97.

Munro, M. 2018. ‘Dimensions of pronunciation’. In O. Kang, R. Thomson and J. Murphy (eds) *The Routledge Handbook of Contemporary English Pronunciation*. New York: Routledge. 413-431.

Narayanan, S. and D. Wang. 2005. ‘Speech rate estimation via temporal correlation and selected sub-band correlation’. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing.* I-413.

Neel, A. 2008. ‘Vowel space characteristics and vowel identification accuracy’. *Journal of Speech, Language, and Hearing Research 51*(3): 574-585.

Nelson, D. 1992. ‘The Foreigner Talk of a Family Physician: An Observational Study’.

Newman, D. 2002. ‘The phonetic status of Arabic within the world's languages: the uniqueness of the lughat al-daad’. *Antwerp Papers in Linguistics 100*: 65-75.

Nissen, S. and R. Fox. 2009. ‘Acoustic and spectral patterns in young children’s stop consonant productions’. *The Journal of the Acoustical Society of America 126*(3): 1369-1378.

Norlin, K. 1987. ‘A phonetic study of emphasis and vowels in Egyptian Arabic’. *Journal ISSN 280*: 526X.

Ohala, J.1994. ‘Acoustic study of clear speech: A test of the contrastive hypothesis’. In *Proceedings of the International Symposium on Prosody*. Yokohama, Japan.75-89.

Olmsted, D. 1951. ‘The phonemes of Yoruba’. *Word 7*(3): 245-249.

Owolabi, D. 2012. ‘Production and perception problems of English dental fricatives by Yoruba speakers of English as a second language’. *Theory and Practice in Language Studies 2*(6): 1108-1113.

Owren, M. 2008. ‘GSU Praat Tools: Scripts for modifying and analyzing sounds using Praat acoustics software’. *Behavior research methods 40*(3): 822-829.

Oyama, S. 1976. ‘A sensitive period for the acquisition of a nonnative phonological system’. *Journal of Psycholinguistic Research 5*(3): 261-283.

Papoušek, H. and M. Papoušek. 1987. ‘Intuitive parenting: a dialectic counterpart to the infant's integrative competence’. In J. Osofsky (ed.) *Handbook of Infant Development (Wiley Series on Personality Processes)*. John Wiley & Sons. 669–720

Patkowski, M. 1989. ‘Age and accent in a second language: A reply to James Emil Flege’. *Applied Linguistics* 11: 73-89.

Pegg, J., J. Werker and P. McLeod. 1992. ‘Preference for Infant-directed over adult-directed speech: Evidence from 7-week-old infants’. *Infant Behavior and Development 15*: 325–345.

Perdue, C. 1984. *Second Language Acquisition by Adult Immigrants: A Field Manual*. Rowley: Newbury House Publishers.

Perdue, C. 1993. *Adult Language Acquisition: Volume 2, the Results: Cross-Linguistic Perspectives* (Vol. 2). Cambridge: Cambridge University Press.

Peterson, G. and H. Barney. 1952. ‘Control methods used in a study of the vowels’. *The Journal of the Acoustical Society of America 24*(2): 175-184.

Peust, C. 1996. ‘Sum: th-substitution’. *The Linguist List 7*: 1108-1110.

Pica, T. 1983. ‘Adult acquisition of English as a second language under different conditions of exposure’. *Language Learning 33*(4): 465-497.

Pica, T., C. Doughty and R. Young. 1986. ‘Making input comprehensible’. *ITL-International Journal of Applied Linguistics 72*(1): 1-25.

Pica, T. and C. Doughty. 1985. ‘The role of group work in classroom second language acquisition’. *Studies in Second Language Acquisition 7*(2): 233-248.

Picheny, M., N. Durlach and L. Braida. 1985. ‘Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech’. *Journal of Speech, Language, and Hearing Research 28*(1): 96-103.

Picheny, M., N. Durlach and L. Braida. 1986. 'Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech'. *Journal of Speech, Language, and Hearing Research 29*(4): 434-446.

Picheny, M., N. Durlach and L. Braida.1989. 'Speaking clearly for the hard of hearing III: An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech'. *Journal of Speech, Language, and Hearing Research 32*(3): 600-603.

Pinker, S. 1979. 'Formal models of language learning'. *Cognition 7*(3): 217-283.

Piske, T., I. MacKay and J. Flege. 2001. 'Factors affecting degree of foreign accent in an L2: A review'. *Journal of Phonetics 29*(2): 191-215.

Piske, T. and M. Young-Scholten. 'Introduction'. In T. Piske and M. Young-Scholten (eds) *Input matters in SLA*. Bristol: Multilingual Matters. 1-26.

Pollack, I. and D. Pisoni. 1971. 'On the comparison between identification and discrimination tests in speech perception'. *Psychonomic Science 24*(6): 299-300.

Polomé, E. 1967. Swahili Language Handbook. New York: The Center for Applied Linguistics.

Prathosh, A., A. Ramakrishnan and T. Ananthapadmanabha. 2015. 'Classification of place-of-articulation of stop consonants using temporal analysis'. In *Sixteenth Annual Conference of the International Speech Communication Association*.

Rajab, H., 2013. 'Developing speaking and writing skills of L1 Arabic EFL learners through teaching of IPA phonetic codes'. *Theory and Practice in Language Studies 3*(4): 653-659.

Rashel, M., 2012. 'Standard Colloquial Bengali and Chatkhil Dialect: a comparative phonological study'. *Language in India: Strength for Today and Bright Hope for Tomorrow 12*: 77-92.

Ratner, N.2000. 'Elicited imitation and other methods for the analysis of trade-offs between speech and language skills in children'. *Methods for studying language production*: 291-312.

Rau, D., H. Chang and E. Tarone. 2009. 'Think or sink: Chinese learners' acquisition of the English voiceless interdental fricative'. *Language Learning 59*(3): 581-621.

Read, C., Y. Zhang, H. Nie, B. Ding. 1986. 'The ability to manipulate speech sounds depends on knowing alphabetic spelling'. *Cognition* 24: 31–44.

Rice, K. 2007. 'Markedness in phonology'. In P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. New York: Cambridge University Press. 79-97.

Richards, B. and C. Gallaway.1993. 'Language acquisition in children: input and interaction'. In R. Asher and J. Simpson (eds) *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon Press.

Rodriguez-Cuadrado, S., C. Baus and A. Costa. 2018. 'Foreigner talk through word reduction in native/non-native spoken interactions'. *Bilingualism: Language and cognition 21*(2): 419-426.

Rosner, B. and J. Pickering. 1994. *Vowel Perception and Production*. Oxford: Oxford University Press.

Routman, R. 2002. 'Teacher talk'. *Educational Leadership 59*(6): 32-35.

Rūķe-Draviņa, V. 1977. 'Modifications of speech addressed to young children in Latvian'. In C. Snow and C. Ferguson (eds) *Talking to Children*. Cambridge: Cambridge University Press.

Ryan, E., J. Hamilton and S. See. 1994. 'Patronizing the old: How do younger and older adults respond to baby talk in the nursing home?'. *International Journal of Aging and Human Development*, 39(1): 21-32.

Saadah, E. 2011. 'The production of Arabic vowels by English L2 learners and heritage speakers of Arabic'. Unpublished Ph.D Dissertation, University of Illinois at Urbana-Champaign.

Sailaja, P. 1999. 'Syllable structure of Telugu'. In *Proceedings of the Fourteenth International Congress of Phonetic Sciences* 4: 743-746.

Salib, M. 1981. *Spoken Arabic of Cairo*. Cairo: AUC Press.

Scarborough, R., O. Dmitrieva, L. Hall-Lew, Y. Zhao, and J. Brenier. 2007. 'An acoustic study of real and imagined foreigner-directed speech'. *Journal of the Acoustical Society of America 121*(5): 3044.

Scarcella, R. and C. Higa. 1982. 'Input and age differences in second language acquisition'. In S. Krashen, R. Scarcella and M. Long (eds) *Child-Adult Differences in Second Language Acquisition*. Rowley, Mass: Newbury House.

Schachter, P. and L. Reid. 2008. 'Tagalog'. In B. Comrie (ed.) *The World's Major Languages*. London: Routledge. 833-855.

Schilling, N. 2013. *Sociolinguistic Fieldwork*. Cambridge: Cambridge University Press.

Schinke-Llano, L. 1983. 'Foreigner talk in content classrooms'. In S. Herbert and L. Michael (eds) *Classroom-oriented Research in Second Language Acquisition*. Rowley: Newbury House Publishers, Inc. 146-165.

Schmidt, R. 1990. 'The role of consciousness in second language learning'. *Applied Linguistics 11*, 129-158.

Schumann, J. 1978. *The Pidginization Process: A model for Second Language Acquisition.* Rowley, Mass: Newbury House Publishers.

Schielzeth, H. 2010. 'Simple means to improve the interpretability of regression coefficients'. *Methods in Ecology and Evolution 1*(2): 103-113.

Scovel, T. 1988. *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech'*. Rowley: Newbury House Publishers.

Segal, J. and R. Newman. 2015. 'Infant preferences for structural and prosodic properties of infant directed speech in the second year of life'. *Infancy 20*(3):339-351.

Seidl, A., A. Cristià, A. Bernard and K. Onishi. 2009. Allophonic and phonemic contrasts in infants' learning of sound patterns. *Language Learning and Development 5*(3): 191-202.

Seliger, H. 1975. 'Maturational constraints in the acquisition of second language accent'. *Language Sciences 36*: 20-22.

Selkirk, E. 1982. 'The syllable'. *The Structure of Phonological Representations 2*: 337-383.

Shehata, A. 2015. 'Problematic Arabic consonants for native English speakers: learners' perspectives'. *The International Journal of Educational Investigations 2*(9): 24-47.

Shinohara, Y. and P. Iverson. 2018. 'High variability identification and discrimination training for Japanese speakers learning English/r/–/l'. *Journal of Phonetics 66*: 242-251.

Showalter, C. and R. Hayes-Harb. 2013. 'Unfamiliar orthographic information and second language word learning: A novel lexicon study'. *Second Language Research 29*(2): 185-200.

Skehan, P. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Smart, J. 1990. 'Pidginization in Gulf Arabic: A first report'. *Anthropological Linguistics:* 83-119.

Smith, J. 2001. 'Lexical category and phonological contrast'. In R. Kirchner, J. Pater, and W. Wikely (eds) *Papers in Experimental and Theoretical Linguistics 6: Workshop on the Lexicon in Phonetics and Phonology.* Edmonton: University of Alberta. 61-72.

Smiljanić, R. and A. Bradlow. 2009. 'Speaking and hearing clearly: Talker and listener factors in speaking style changes'. *Language and Linguistics Compass 3*(1): 236-264.

Snidecor, J., L. Malbry and E. Hearsey.1944. 'Methods of training talkers for increasing intelligibility in noise'. *ORSD Report 3178. New York: The Psychological Corporation*.

Snow, C. 1972. 'Mothers' speech to children learning language'. *Child Development*: 549-565.

Snow, C., R. Eeden and P. Muysken. 1981. 'The interactional origins of foreigner talk: Municipal employees and foreign workers'. *International Journal of the Sociology of Language 1981*(28) 81-92.

Snow, C. and M. Hoefnagel-Höhle. 1982. 'School-age second language learners' access to simplified linguistic input'. *Language Learning 32*(2): 411-430.

Soderberg, C. and K. Olson. 2008. 'Indonesian'. *Journal of the International Phonetic Association 38*(2): 209-213.

Song, J., K. Demuth and J. Morgan. 2010. ‘Effects of the acoustic properties of infant-directed speech on infant word recognition’. *The Journal of the Acoustical Society of America 128*(1): 389-400.

Stevens, P. 1960. ‘Spectra of fricative noise in human speech.’ Lang. Speech 3: 32–49.

Stevens, K. and S. Blumstein. 1978. ‘Invariant cues for place of articulation in stop consonants’. *The Journal of the Acoustical Society of America 64*(5): 1358-1368.

Stevens, K. 1980. ‘Acoustic correlates of some phonetic categories’. *The Journal of the Acoustical Society of America 68*(3): 836-842.

Stevens, K. 1994. *Acoustic Phonetics*. Massachusetts: MIT Press.

Strange, W. 1987. ‘Information for vowels in formant transitions’. *Journal of Memory and Language 26*(5): 550-557.

Strange, W. and V. Shafer. 2008. ‘Speech perception in second language learners’. In J. Hansen and M. Zampini (eds) *Phonology and Second Language Acquisition.* Philadelphia*:* John Benjamins. 153-191.

Strange, W. and V. Shafer. 2008. ‘Speech perception in second language learners: The re-education of selective perception’. *Phonology and Second Language Acquisition 36*: 153-192.

Streeter, L., R. Krauss, W. Apple and N. Macdonald. 1978. ‘Acoustic consequences and perceptual indicators of stress’. *The Journal of the Acoustical Society of America 64*(S1): S115-S115.

Sundberg, U. 1998. ‘Mother tongue-phonetic aspects of infant-directed speech’. Ph.D Dissertation, University of Stockholm.

Sundberg, U. and F. Lacerda. 1999. ‘Voice onset time in speech to infants and adults’. *Phonetica 56*(3-4): 186-199.

Sundberg, U. 2001. ‘Consonant specification in infant-directed speech. Some preliminary results from a study of Voice Onset Time in speech to one-year-olds’. *Lund University Department of Linguistics Working Papers 49*: 148-151.

Supriyadi, T. and J. Julia. 2019. ‘The problem of students in reading the Quran: A reflective-critical treatment through Action research’. *International Journal of Instruction 12*(1): 311-326.

Suyanto, S., S. Hartati, A. Harjoko and D. Van Compernolle. 2016. ‘Indonesian syllabification using a pseudo nearest neighbour rule and phonotactic knowledge’. *Speech Communication 85*: 109-118.

Swanson, L., L. Leonard and J. Gandour. 1992. ‘Vowel duration in mothers' speech to young children. *Journal of Speech, Language, and Hearing Research 35*(3): 617-625.

Swanson, L. and L. Leonard. 1994. ‘Duration of function-word vowels in mothers' speech to young children’. *Journal of Speech, Language, and Hearing Research 37*(6): 1394-1405.

Symonds, M. and A. Moussalli. 2011. ‘A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike’s information criterion’. *Behavioral Ecology and Sociobiology 65*(1): 13-21.

Synnestvedt, A. 2010. ‘Voice onset time in infant-directed speech at two ages’. Unpublished Doctoral Dissertation, University of Maryland.

Tabain, M. 2001. ‘Variability in fricative production and spectra: Implications for the hyper-and hypo-and quantal theories of speech production’. *Language and Speech 44*(1): 57-93.

Taqi, H. 2010. ‘Two ethnicities, three generations: phonological variation and change in Kuwait’. Unpublished PhD Dissertation, Newcastle University.

Tarone, E. 1980. ‘Communication strategies, foreigner talk, and repair in interlanguage 1’. *Language Learning 30*(2): 417-428.

Tarone, E. and M. Bigelow. 2005. ‘Impact of literacy on oral language processing: Implications for second language acquisition research’. *Annual Review of Applied Linguistics 25*: 77-97.

Tarone, E., M. Bigelow and K. Hansen. 2009. *Literacy and Second Language Oracy-Oxford Applied Linguistics.* Oxford: Oxford University Press.

Tarone, E. 2010. ‘Second language acquisition by low-literate learners: An under-studied population’. *Language Teaching 43*(1): 75-83.

Team, R.C. 2017. ‘R: a language and environment for statistical computing’. R Foundation for Statistical Computing. Austria, Vienna.

Templin, M. 1957. ‘Certain language skills in children; their development and interrelationships’. *The Institute of Child Welfare Monographs*. Minneapolis: University of Minnesota.

Thompson, H. 2012. *Bengali*. Amsterdam: John Benjamins Publishing.

Thompson, I. 1991. ‘Foreign accents revisited: The English pronunciation of Russian immigrants’. *Language Learning 41*(2): 177-204.

Tolhurst, G. 1957. ‘Effects of duration and articulation changes on intelligibility, word reception and listener preference’. *Journal of Speech and Hearing Disorders 22*(3): 328-334.

Traunmüller, H., 1990. ‘A note on hidden factors in vowel perception experiments’. *The Journal of the Acoustical Society of America 88*(4): 2015-2019.

Trofimovich, P. and W. Baker. 2006. ‘Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech’. *Studies in Second Language Acquisition 28*(01): 1-30.

Uchanski, R. 2005. ‘Clear speech’. In D. Pisoni and R. Remez (eds) *The Handbook of Speech Perception*. Malden: Blackwell Publishing Ltd. 207-235.

ud Dowla Khan, S. 2010. ‘Bengali (Bangladeshi Standard)’. *Journal of the International Phonetic Association*: 221-225.

Uther, M., M. Knoll and D. Burnham. 2007. ‘Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech’. *Speech Communication 49*(1): 2-7.

Valdman, A. 1981. ‘Sociolinguistic aspects of foreigner talk’. *International Journal of the Sociology of Language 1981*(28): 41-52.

Van Engen, K., M. Baese-Berk, R. Baker, A. Choi, M. Kim and A. Bradlow. 2010. ‘The Wildcat corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles’. *Language and Speech* 53 (4): 510-540.

Varonis, E. and S. Gass. 1982. ‘The comprehensibility of non-native speech’. *Studies in Second Language Acquisition 4*(2): 114-136.

Vennemann, T. 1988. *Preference Laws for Syllable Structure*. Berlin: Walter de Gruyter.

Wang, Y., D. Houston and a. Seidl. 2019. ‘Acoustic properties of infant-directed speech’. In S. Fruhholz and P. Belin (eds.) *The Oxford Handbook of Voice Perception*. Oxford: Oxford University Press, USA. 93-116.

Wasala, A. and K. Gamage. 2005. ‘Research report on phonetics and phonology of Sinhala’. *Language Technology Research Laboratory, University of Colombo School of Computing 35*.

Watson, J. 2002. *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.

Watson, J. 2011. ‘Arabic dialects (general article)’. In S. Weninger, G. Khan, M. Streck and J Watson (eds) *The Semitic Languages: An International Handbook*. Handbooks of Linguistics and Communication Science. Berlin: Walter de Gruyter. 851-896.

Wellman, B., I. Case, I. Mengert, and D. Bradbury. 1931. ‘Speech sounds of young children’. University of Iowa Studies: Child Welfare.

Werker, J. and R. Tees. 1984. ‘Cross-language speech perception: Evidence for perceptual reorganization during the first year of life’. *Infant Behavior and Development 7*(1): 49-63.

Werker, J. and J., Logan. 1985. ‘Cross-language evidence for three factors in speech perception’. *Attention, Perception & Psychophysics 37*(1): 35-44.

Werker, J. 1995. ‘Exploring developmental changes in cross-language speech perception’. *Language: An Invitation to Cognitive Science 1*: 87-106.

Werker, J. and R. Tees. 1999. ‘Influences on infant speech processing: Toward a new synthesis’. *Annual Review of Psychology 50*(1): 509-535.

Wesche, M. 1994. ‘Input and interaction in second language acquisition’. In C. Gallaway and B. Richards (eds) *Input and Interaction Language Acquisition*. Cambridge: Cambridge University Press. 219-249.

Winter, B. 2013. ‘Linear models and linear mixed effects models in R with linguistic applications’. arXiv:1308.5499.

Yeou, M. 1997. ‘Locus equations and the degree of coarticulation of Arabic consonants’. *Phonetica* 54: 187–202.

Young-Scholten, M. 1995. ‘The negative effects of ‘positive’evidence’. In L. Eubank, L. Selinker and M. Smith (eds) *The Current State of Interlanguage: Studies in Honor of William E. Rutherford*. Amsterdam: John Benjamins Publishing. 107.

Young-Scholten, M. and Archibald, J., 2000. Second language syllable structure. In J. Archibald (ed.) *Second Language Acquisition and Linguistic Theory*. Oxford: Wiley-Blackwell. 64-101.

Young-Scholten, M. and N. Strom. 2006. ‘First-time L2 readers: Is there a critical period?’. *LOT Occasional Series 6*: 45-68.

Young-Scholten, M. 2013. ‘Low-educated immigrants and the social relevance of second language acquisition research’. *Second Language Research 29*(4): 441-454.

Young-Scholten, M. and M. Langer. 2015. ‘The role of orthographic input in second language German: Evidence from naturalistic adult learners’ production’. *Applied Psycholinguistics 36*(1): 93-114.

Young-Scholten, M. and R. Naeb. 2017.’ International training of teachers of low-educated adult migrants’. In J. Beacco, H. Krumm, D. Little and P.Thalgott (eds) *The Linguistic Integration of Adult Migrants: Some Lessons from Research*. De Gruyter. 419-424.

Youssef, I. 2013. ‘Place assimilation in Arabic: Contrasts, features, and constraints’. Unpublished Ph.D. Dissertation, University of Tromsø.

Zawaydeh, B. 1998. ‘Gradient uvularization spread in Ammani-Jordanian Arabic’. In E. Benmamoun, M. Eid and N. Haeri (eds) *Perspectives on Arabic Linguistics XI: Papers from The Eleventh Annual Symposium On Arabic Linguistics.* Amsterdam: John Benjamins Publishing. 117-41.

Zawaydeh, B. 1999. ‘The phonetics and phonology of gutturals in Arabic’. Unpublished Ph.D Dissertation, Indiana University, Bloomington.

Zec, D. 2007. ‘The syllable’. In P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press. 161-194.

Zhao, S. 2010. ‘Stop-like modification of the dental fricative/ð/: An acoustic analysis’. *The Journal of the Acoustical Society of America 128*(4): 2009-2020.

Zuengler, J., 1991. ‘Accommodation in native-nonnative interactions: Going beyond the “what” to the “why” in second language research’. In H. Giles and J. Coupland (eds) *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Cambridge: Cambridge University Press. 223-244.

Zuraida, I. and N. Fitr. 2019. ‘Foreigner Talk in Conversation from YouTube Channel Abroad in Japan’. *English Journal Literacy Utama 3*(1): 56-65.

Appendix 1. Participant information sheet and consent form (Employer)

PARTICIPANT INFORMATION SHEET (Employer)

Title of project: Effects of Input, Experience and Native Language on the Production of Arabic Sounds by Foreign Domestic Helpers in a Naturalistic Setting

Name of supervisor: Dr. Ghada Khattab and Prof. Martha Young-Scholten

Email: martha.young-scholten@ncl.ac.uk

Work phone number: +44 (0) 191 208 7751

Name of researcher: Azza Al-Kendi

Email: A.N.S.Al-kendi2@newcastle.ac.uk or kindisoul@gmail.com Mobile: +447492777626; +447492807807

Contact address: School of English Literature language and Linguistics, Percy Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.

You are invited to participate in a project on the above title by the above mentioned researcher. Before you decide to take part in the project, you need to understand some basic information on why the research is being conducted. Please take your time to read the information on this form before deciding to participate in the project and do feel free to ask any questions to clarify the information provided.

Purpose and aims of the research

The purpose of this research is to examine how adult non-native speakers acquire a second language in a naturalistic setting (with no formal instruction). Research of this type is essential because the findings could recommend alternative methods of teaching oral foreign languages to non-native speakers.

Participation selection

You have been approached to participate in this project because you have a domestic helper with whom you engage in interactions.

Voluntary participation

If you choose to participate, you will be asked to sign two consent forms; one to indicate your willingness to participate and the other to confirm that the domestic helper could spare some of her time to participate in the study as well as make sure her participation is voluntary. Your participation is voluntary. You have the right to withdraw from the study at any time without any Consequences. You only need to inform the researcher beforehand. If you withdraw, your data will only be used if you authorize the researcher to do so. Otherwise it will be destroyed.

What is involved in participating

If you agree to participate in this project, you will be asked to sit for a game with the domestic helper for about 15-20 minutes. During this game, you and the domestic helper will be given two pictures with identical scenes but that differ in some aspects. You will be asked to negotiate with the domestic helper the differences between the two pictures. Afterwards, the researcher will ask the domestic helper alone to do another task. You will be asked to leave the room because your presence might affect the way the domestic helper would interact with the researcher and this could affect the domestic helper's responses. The game task will be recorded using a digital recorder.

Benefits and risk

Participation does not involve any known or anticipated risk for you. However, participation may cause inconvenience, as it will require about one hour of your time.

Anonymity and confidentiality

Your real name will never be used neither in written nor verbal reports of this research. Instead, pseudonym or code will be used.

Confidentiality, storage and usage of data

Your confidentiality and that of the data will be protected during and after the research. The recordings and other electronic documents will be stored in password-secured server. Hard copies of transcriptions and other information documents will be stored in a locked cabinet accessible only to the researcher.

معلومات للمشارك

عنوان الدراسه: إكتساب العمال الأجانب للغه العربيه في بيئه غير تعليميه

إسم المشرف: بروفيسوره مارثا يونج شولتن

الإيميل:ku.ca.lcn@netlohcs-gnuoy.ahtram

رقم هاتف العمل: +441912087751

إسم الباحث: عزة الكندي

الإيميل:ku.ca.lcn@2idnek-lA.S.N.A أو moc.liamg@luosidnik

رقم الهاتف: +447492777626

العنوان للتواصل:

School of English Literature, language and Linguistics, Percy Building, Newcastle University, Newcastle Upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom:

أنت مدعو للمشاركه في مشروع بالعنوان المذكور سابقا ومن قبل الباحث المذكور أعلاه. قبل الموافقه على المشاركه عليك الإطلاع على بعض المعلومات عن سبب إقامة هذا المشروع. نرجو منك قراءة المعلومات في هذا الملف قبل اتخاذ قرار المشاركه والتقدم بأي أسئله أو استفسارات لتوضيح أي معلومه.

هدف المشروع

الهدف من هذا المشروع هو دراسة كيفية اكتساب الشخص البالغ للغه أخرى في بيئه غير تعليميه وبدون الحصول على أي تعليم. هذه الدراسه مهمه لأنها سوف تحصل على نتائج من الممكن الإستفاده منها في تعليم اللغات الأخرى لغير المتحدثين بها.

إختيار المشاركين

لقد تم اختيارك للمشاركه في هذه الدراسه لأن لديك عاملة منزل تقوم بالتحدث معها من حين لآخر.

المشاركه التطوعيه

إذا وافقت على المشاركه, سوف يطلب منك التوقيع على إستمارتي موافقه. الأولى للإقرار برغبتك في المشاركه. والثانيه لتأكيد مقدرة العامله على إعطاء جزء من وقت فراغها للمشاركه والـاكيد بأن مشاركتها تطوعيه. مشاركتك في هذا المشروع تطوعيه ولديك الحق في الإنسحاب من المشروع في أي وقت بدون التعرض لأي مسائله. إذا قررت الإنسحاب سوف يقوم الباحث باستخدام بياناتك بعد موافقتك وإلا فسوف يقوم بالتخلص منها.

محتوى المشاركه

اذا وفقت على المشاركه سوف يطلب منك الإنضمام مع عاملة المنزل في لعبة تستغرق بين 15-20 دقيقه. سوف يقوم الباحث بإعطائك وإعطاء عاملة المنزل صورتين متشابهتان ولكن فيهما بعض الإختلافات. مهمتك أن تقوم بإكتشاف هذه الاختلافات وذلك بالتحاور مع عاملة المنزل. سوف يتم تسجيل صوتك في هذه المشاركه من خلال مسجل صوتي. سوف يطلب منك بعد ذلك المغادره لأستكمال التمرين الآخر مع العامله فقط وذلك لأن وجودك قد يؤثر على طريقة تجاوب العامله مع الباحث مما قد يؤثر على إجاباتها.

<u>المنافع والأخطار</u>

المشاركه في هذا المشروع لن تتسبب لك بأي مخاطر تذكر. قد تستغرق التمارين ما يقارب الساعه الواحده وهذا قد يأخذ جزءا من وقتك ووقت العامله.

<u>الخصوصيه والسريه</u>

سوف يتم المحافظه على خصوصية مشاركتك في المشروع. التسجيلات وغيرها من الملفات الإلكترونيه سوف يتم حفظها في كمبيوتر آمن بكلمة سر. الملفات الغير إلكترونيه ستحفظ في خزانه مغلقه وستتوفر فقط للباحث والمشرف.

<u>الخصوصيه, تخزين البيانات واستخدامها</u>

لن يتم استخدام اسمك في أي تقرير كتابي أو شفهي في هذا المشروع. سوف نقوم باستخدام اسم مستعار أو رمز معين للإشاره للمشارك. أي اسم ستقوم بذكره في المحادثه سيتم تغييره خلال عملية تحليل النتائج وحذفه من التسجيلات.

<u>نشر النتائج</u>

نتائج هذه الدراسه ستسخدم في مشروع الباحث. قد تستخدم البيانات والنتائج أيضا في بحوث ومنشورات ومؤتمرات خارجيه.

<u>معلومات إضافيه</u>

إذا كان لديك أية أسئله او إستفسارات عن هذا المشروع أو ترغب في الحصول على المزيد من المعلومات, نرجو منك التواصل مع الباحث أو المشرف بطرق التواصل الموفره أعلاه.

CONSENT FORM (Employer)

Name of supervisor: Prof. Martha Young-Scholten & Dr. Ghada Khattab

Email: martha.young-scholten@ncl.ac.uk

Work phone number: +44 (0) 191 208 7751

Name of researcher: Azza Al-Kendi

Email: A.N.S.Al-kendi2@newcastle.ac.uk or kindisoul@gmail.com Mobile: +447492777626; +447492807807

Contact address: School of English Literature language and Linguistics, Percy Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.

I, the undersigned participant confirms the following:

My domestic helper could spare some of her time (around one hour) to participate in the study.

Her participation is fully voluntary and she has the right to withdraw from the study at any time.

I am aware that she could refuse to take part in the study and that she would not be coerced to participate.

Her participation will not have negative effects for any of us.

Name of participant giving consent ---

Signature of participant ---

Name of researcher taking consent ---

Signature of researcher ---

Date ---

استمارة موافقه (ربة المنزل)

عنوان الدراسه: إكتساب العمال الأجانب للغه العربيه في بيئه غير تعليميه

إسم المشرف: بروفيسوره مارثا يونج شولتن

الإيميل:ku.ca.lcn@netlohcs-gnuoy.ahtram

رقم هاتف العمل: +441912087751

إسم الباحث: عزة الكندي

الإيميل:ku.ca.lcn@2idnek-lA.S.N.A أو moc.liamg@luosidnik

رقم الهاتف: +447492777626

العنوان للتواصل:

School of English Literature, language and Linguistics, Percy Building, Newcastle University, Newcastle Upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom:

أنا الموقع أدناه أقر بالآتي:

عاملة المنزل لدي تستطيع أن تخصص جزء من وقتها ما يقارب الساعه للمشاركه في الدراسة.

مشاركتها تطوعية وتستطيع الانسحاب من الدراسة في أي وقت تشاء.

أعي بأن لديا الحق في رفض المشاركة وانها لن تجبر على المشاركه ضد رغبتها.

مشاركتها لن تؤثر بطريقة سلبيه على أي منا.

اسم المشارك ------------------------ التاريخ----------------- التوقيع-------------------------

الباحث -------------------------- التاريخ----------------- التوقيع--------------------------

Appendix 2. Participant information sheet and consent form (Foreign Domestic Helper)

PARTICIPANT INFORMATION SHEET (Foreign Domestic Helper)

Title of project: Effects of Input, Experience and Native Language on the Production of Arabic Sounds by Foreign Domestic Helpers in a Naturalistic Setting

Name of supervisor: Dr. Ghada Khattab and Prof. Martha Young-Scholten

Email: martha.young-scholten@ncl.ac.uk

Work phone number: +44 (0) 191 208 7751

Name of researcher: Azza Al-Kendi

Email: A.N.S.Al-kendi2@newcastle.ac.uk or kindisoul@gmail.com Mobile: +447492777626; +447492807807

Contact address: School of English Literature language and Linguistics, Percy Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.

You are invited to participate in a project on the above title by the above mentioned researcher. Before you decide to take part in the project, you need to understand some basic information on why the research is being conducted. Please take your time to read the information on this form before deciding to participate in the project and do feel free to ask any questions to clarify the information provided.

Purpose and aims of the research

The purpose of this research is to examine how adult non-native speakers acquire a second language in a naturalistic setting (with no formal instruction). Research of this type is essential because the findings could recommend alternative methods of teaching oral foreign languages to non-native speakers.

Participation selection

You have been approached to participate in this project because you are a non-native speaker of Arabic who is learning the language in a naturalistic setting.

Voluntary participation

If you choose to participate, you will be asked to sign a consent form to indicate your willingness to participate. Your participation is voluntary. You have the right to withdraw from the study at any time without any consequences. You only need to inform the researcher beforehand. If you withdraw, your data will only be used if you authorize the researcher to do so. Otherwise it will be destroyed.

What is involved in participating

If you agree to participate in this project, you will be asked to sit for a game with your employer for about 15-20 minutes. During this game, you and your employer will be given two pictures with identical scenes but that differ in some aspects. You will be asked to negotiate with your emploeyr the differences between the two pictures. Afterwards, the researcher will ask you alone to do another task. Your employer will be asked to leave the room because her presence might affect the way you would interact with the researcher and this could affect your responses. The game task will be recorded using a digital recorder.

Benefits and risk

Participation does not involve any known or anticipated risk for you. However, participation may cause inconvenience, as it will take about one hour of your time.

Anonymity and confidentiality

Your real name will never be used neither in written nor verbal reports of this research. Instead, pseudonym or code will be used.

Confidentiality, storage and usage of data

Your confidentiality and that of the data will be protected during and after the research. The recordings and other electronic documents will be stored in password-secured server. Hard copies of transcriptions and other information documents will be stored in a locked cabinet accessible only to the researcher.

معلومات للمشارك

عنوان الدراسه: إكتساب العمال الأجانب للغه العربيه في بيئه غير تعليميه

إسم المشرف: بروفيسوره مارثا يونج شولتن

الإيميل:ku.ca.lcn@netlohcs-gnuoy.ahtram

رقم هاتف العمل: +441912087751

إسم الباحث: عزة الكندي

الإيميل:ku.ca.lcn@2idnek-lA.S.N.A أو moc.liamg@luosidnik

رقم الهاتف: +447492777626

العنوان للتواصل:

School of English Literature, language and Linguistics, Percy Building, Newcastle University, Newcastle Upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom:

أنت مدعو للمشاركه في مشروع بالعنوان المذكور سابقا ومن قبل الباحث المذكور أعلاه. قبل الموافقه على المشاركه عليك الإطلاع على بعض المعلومات عن سبب إقامة هذا المشروع. نرجو منك قراءة المعلومات في هذا الملف قبل اتخاذ قرار المشاركه والتقدم بأي أسئله أو استفسارات لتوضيح أي معلومه.

هدف المشروع

الهدف من هذا المشروع هو دراسة كيفية اكتساب الشخص البالغ للغه أخرى في بيئه غير تعليميه وبدون الحصول على أي تعليم. هذه الدراسه مهمه لأنها سوف تحصل على نتائج من الممكن الإستفاده منها في تعليم اللغات الأخرى لغير المتحدثين بها.

إختيار المشاركين

لقد تم اختيارك للمشاركه في هذه الدراسه لأنك متحدث باللغة العربيه كلغة اخرى اكتسبتها في ظروف طبيعية بالتحاور مع أفراد المنزل.

المشاركه التطوعيه

إذا وافقت على المشاركه, سوف يطلب منك التوقيع على إستمارة موافقه للإقرار برغبتك في المشاركه. مشاركتك في هذا المشروع تطوعيه ولديك الحق في الإنسحاب من المشروع في أي وقت بدون التعرض لأي مسائله. إذا قررت الإنسحاب سوف يقوم الباحث باستخدام بياناتك بعد موافقتك وإلا فسوف يقوم بالتخلص منها.

محتوى المشاركه

اذا وفقت على المشاركه سوف يطلب منك الإنضمام مع ربة المنزل في لعبة تستغرق بين 15-20 دقيقه. سوف يقوم الباحث بإعطائك وإعطاء ربة المنزل صورتين متشابهتان ولكن فيهما بعض الإختلافات. مهمتكما أن تقوما بإكتشاف هذه الاختلافات وذلك بالتحاور. سوف يطلب من ربة المنزل بعد ذلك مغاذرة المكان ليقوم الباحث بإجراء تمرين آخر لك وحدك. هذا التمرين عبارة عن صور سيقوم الباحث بعرضها لك ومهمتك ان تقوم بتسمية هذه الصور. سوف يتم تسجيل صوتك في هذه المشاركه من خلال مسجل صوتي.

المنافع والأخطار

المشاركه في هذا المشروع لن تتسبب لك بأي مخاطر تذكر. قد تستغرق التمارين ما يقارب الساعه الواحده وهذا قد يأخذ جزءا من وقتك..

الخصوصيه والسريه

سوف يتم المحافظه على خصوصية مشاركتك في المشروع. التسجيلات وغيرها من الملفات الإلكترونيه سوف يتم حفظها في كمبيوتر آمن بكلمة سر. الملفات الغير إلكترونيه ستحفظ في خزانه مغلقه وستتوفر فقط للباحث والمشرف.

الخصوصيه, تخزين البيانات واستخدامها

لن يتم استخدام اسمك في أي تقرير كتابي أو شفهي في هذا المشروع. سوف نقوم باستخدام اسم مستعار أو رمز معين للإشاره للمشارك. أي اسم ستقوم بذكره في المحادثه سيتم تغييره خلال عملية تحليل النتائج وحذفه من التسجيلات.

نشر النتائج

نتائج هذه الدراسه ستسخدم في مشروع الباحث. قد تستخدم البيانات والنتائج أيضا في بحوث ومنشورات ومؤتمرات خارجيه.

معلومات إضافيه

إذا كان لديك أية أسئله او إستفسارات عن هذا المشروع أو ترغب في الحصول على المزيد من المعلومات, نرجو منك التواصل مع الباحث أو المشرف بطرق التواصل الموفره أعلاه.

CONSENT FORM

Name of supervisor: Dr. Ghada Khattab and Prof. Martha Young-Scholten

Email: martha.young-scholten@ncl.ac.uk

Work phone number: +44 (0) 191 208 7751

Name of researcher: Azza Al-Kendi

Email: A.N.S.Al-kendi2@newcastle.ac.uk or kindisoul@gmail.com Mobile: +447492777626; +447492807807

Contact address: School of English Literature language and Linguistics, Percy Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.

I, the undersigned participant confirm that (please tick box appropriately):

1	I have read and understood the information about the project as provided on the information sheet.	☐
2	I have been given the opportunity to ask questions about the project and my participation.	☐
3	I agree that my participation is voluntary.	☐
4	I understand that I can withdraw at any time without giving reasons or being penalised nor will I be questioned for withdrawing.	☐
5	I understand that a voice recorder will be used to collect data and I agree to my voice being recorded for the purpose of this research project.	☐
6	The procedures regarding confidentiality and anonymity have been clearly explained to me.	☐
7	I understand that the recording of my voice and other accompanying materials may be stored in password-protected files computers.	☐
8	I understand that anonymised extracts of my data may be used in research, publication and conferences.	☐
9	Storage and usage of data has been explained to me	☐
10	I understand that the researcher will remove my data if I decided to withdraw from the study at any time.	☐

Name of participant giving consent --

Signature of participant --

Name of researcher taking consent --

Signature of researcher --

Date --

إستمارة موافقه

إسم المشرف: بروفيسوره مارثا يونج شولتن
الإيميل:ku.ca.lcn@netlohcs-gnuoy.ahtram
رقم هاتف العمل: +441912087751
إسم الباحث: عزة الكندي
الإيميل:ku.ca.lcn@2idnek-lA.S.N.A أو moc.liamg@luosidnik
رقم الهاتف: +447492777626

العنوان للتواصل:
School of English Literature, language and Linguistics, Percy Building, Newcastle University, Newcastle Upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom:

أقر أنا الموقع أدناه بأنني (ضلل المربع المناسب):

1	أؤكد بأنني قد قرأت معلومات المشارك المذكوره أعلاه.	
2	أؤكد أنني قد أتيحت لي الفرصة لطرح الأسئلة عن المشروع وعن مشاركتي.	
3	أفهم بأن لدي الحق بالإنسحاب في أي وقت بدون ذكر الأسباب أو التعرض لأي مسائله .	
4	أفهم بأن مسجل صوتي سوف يستخدم لتسجيل المحادثه لهذا المشروع.	
5	لقد تم إبلاغي عن سرية هذه المقابله وأن هويتي ستبقى مجهوله.	

6	أفهم بأن جميع التسجيلات والملفات الأخرى ستحفظ في كمبيوتر آمن بكلمة سر.	
7	أفهم بأن محادثاتي المسجله قد تستخدم في بحوث ومنشورات ومؤتمرات.	
8	أوافق بأن مشاركتي تطوعيه.	
9	لقد تم إبلاغي بطريقة حفظ واستخدام المعلومات.	

اسم المشارك ------------------------ التاريخ----------------- التوقيع--------------------------

الباحث -------------------------- التاريخ----------------- التوقيع-------------------------

Appendix 3. Language history questionnaire

Participant code: ------------------------------------ Today's date: -----------------------------------

Part A

1. Age:--------------------------------
2. Country of Origin:-------------------------------------
3. Education (years of schooling or degree obtained):-----------------------------------
4. Age of first exposure to Arabic:--
5. Age of arrival to Oman:--
6. How long have you been in Oman? --
7. Have you worked in any other Arab country before? If yes, where and for how long?

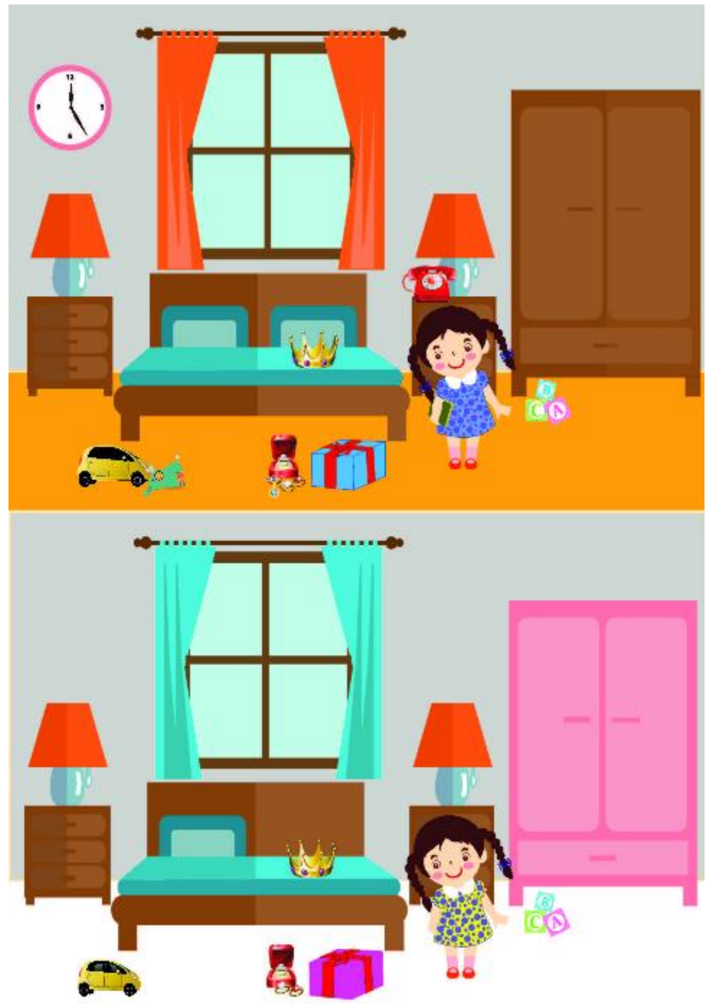
Part B

1. First language:----------------------------------
2. Other languages spoken and proficiency level:---

   ----
3. Literacy in Arabic:

Reading: ------------------------------ Writing: ------------------------------

Appendix 4. Pictures used in spot the difference task

Theme: Bedroom	
NS/Domestic Helper Version	**NS/Native Adult Version**
	

Theme: Living Room	
NS/Domestic Helper Version	**NS/Native Adult Version**
	

<table>
<tr><th colspan="2">Theme: Kitchen</th></tr>
<tr><th>NS/Domestic Helper Version</th><th>NS/Native Adult Version</th></tr>
<tr><td></td><td></td></tr>
</table>

Appendix 5. Stimuli used in the perception experiment before randomization

(the grey shade represents repetitive trials that have been deleted)

Trial / Contrast	Trial 1	Trial 2	Trial 3	Trial 4
/Ɵ/-/ð/	/Ɵan/ /ðan/	/ðan/ /Ɵan/	/Ɵan/ /Ɵan/	/ðan/ /ðan/
	Trial 5	Trial 6	Trial 7	Trial8
/χ/-/ʁ /	/χan/ /ʁan/	/ʁan/ /χan/	/χan/ /χan/	/ʁan/ /ʁan/
	Trial 9	Trial 10	Trial 11	Trial 12
/ħ/-/ʕ/	/ħan/ /ʕan/	/ʕan/ /ħan/	/ħan/ /ħan/	/ʕan/ /ʕan/
	Trial 13	Trial 14	Trial 15	Trial 16
/tˤ /-/t/	/tˤan/ /tan/	/tan/ /tˤan/	/tˤan/ /tˤan/	/tan/ /tan/
	Trial 17	Trial 18	Trial 19	
/ðˤ/-/ð/	/ðˤan/ /ðan/	/ðan/ /ðˤan/	/ðˤan/ /ðˤan/	
	Trial 20	Trial 21	Trial 22	Trial 23
/s/-/sˤ /	/san/ /s̠ ˤ an/	/s ˤ an/ /san/	/sˤ an/ /sˤ an/	/san/ /san/
	Trial 24	Trial 25		Trial 26
/q/-/χ /	/qan/ /χan/	/χan/ /qan/		/qan/ /qan/
	Trial 27	Trial 28		Trial 29
/ħ /-/h/	/ħan/ /han/	/han/ /ħan/		/han/ /han/
	Trial 30	Trial 31	Trial 32	
/k/-/q/	/kan/ /qan/	/qan/ /kan/	/kan/ /kan/	
	Trial 33	Trial 34		
/χ/-/ħ /	/χan/ /ħan/	/ħan/ /χan/		
	Trial 35	Trial 36	Trial 37	Trial 38

/r/-/l/	/ran/ /lan/	/lan/ /ran	/ran/ /ran/	/lan/ /lan/

	Trial 39	Trial 40		Trial 41
/r/-/w/	/ran/ /wan/	/wan/ /ran/		/wan/ /wan/
	Trial 42	Trial 43		
/χ/-/k/	/χan/ /kan/	/kan/ /χan/		
	Trial 44	Trial 45		
/Ɵ/-/s/	/Ɵan/ /san/	/san/ /Ɵan/		
	Trial 46	Trial 47		
/Ɵ/-/t/	/Ɵan/ /tan/	/tan/ /Ɵan/		
	Trial 48	Trial 49		Trial 51
/ð/-/z/	/ðan/ /zan/	/zan/ /ðan/		/zan/ /zan/
	Trial 50	Trial 51		Trial 53
/ð/-/d/	/ðan/ /dan/	/dan/ /ðan/		/dan/ /dan/
	Trial 54	Trial 55		Trial 56
/ʁ/-/g/	/ʁan/ /gan/	/gan/ /ʁan/		/gan/ /gan/

Appendix 6. Stimuli used in the perception experiment after randomization

Trial No.	Test Item	Trial No.	Test Item
1	/Ɵan/ /ðan/	29	/ðˤan/ /ðˤan/
2	/tan/ /tˤan/	30	/gan/ /gan/
3	/san/ /san/	31	/lan/ /ran/
4	/sˤan/ /sˤan/	32	/tˤan/ /tan/
5	/kan/ /χan/	33	/χan/ /ħan/
6	/ðan/ /ðˤan/	34	/ðˤan/ /ðan/
7	/Ɵan/ /Ɵan/	35	/ran/ /lan/
8	/ħan/ /ʕan/	36	/zan/ /zan/
9	/san/ /sˤan/	37	/ħan/ /ħan/
10	/dan/ /dan/	38	/sˤan/ /san/
11	/ðan/ /ðan/	39	/han/ /ħan/
12	/ʁan/ /χan/	40	/gan/ /ʁan/
13	/χan/ /χan/	41	/kan/ /qan/
14	/lan/ /lan/	42	/qan/ /χan/
15	/ðan/ /dan/	43	/ran/ /ran/
16	/ħan/ /han/	44	/ðan/ /Ɵan/
17	/zan/ /ðan/	45	/san/ /Ɵan/
18	/han/ /han/	46	/Ɵan/ /tan/
19	/wan/ /ran/	47	/ʕan/ /ʕan/
20	/ʁan/ /gan/	48	/Ɵan/ /san/
21	/ðan/ /zan/	49	/wan/ /wan/
22	/χan/ /kan/	50	/χan/ /ʁan/
23	/ʕan/ /ħan/	51	/tˤan/ /tˤan/
24	/ħan/ /χan/	52	/qan/ /kan/
25	/kan/ /kan/	53	/χan/ /qan/
26	/qan/ /qan/	54	/ʁan/ /ʁan/
27	/tan/ /Ɵan/	55	/dan/ /ðan/
28	/tan/ /tan/	56	/ran/ /wan/

Appendix 7. AX perception task answer sheet

Trial No.	Same	Different
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		

Name of

Participant………………….Date:………….

Appendix 8. Pictures used in the picture-naming task

19.

20.

21.

22.

23.

24.

25.

26.

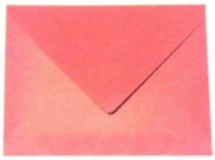
27.

28.

29.

30.

13.



14.

15.

16.

17.

18.

31.

32.

33.

34.

35.

36.

Appendix 9. Sensitivity scores of all subjects

Participant Code	Sensitivity Score
Foreign Domestic Helper Group	
Si1	1.494
Sw2	2.141
Sw4	2.228
Be1	1.044
Ta5	2.384
In1	1.206
Ta4	3.292
Ta3	1.694
Si4	1.238
Ta1	2.141
Si3	2.487
Yo	1.238
Ta2	3.216
Or	3.09
Be2	0
Sw1	4.044
Si2	4.17
Lu	1.227
Sw3	3.369
In2	1.771
Native Speaker Control Group	
Wl	4.126
Mn	3.526
Kw	6.18
Hd	6.18
Lm	6.18
Sb	4.645
Sa	4.735
Gh	4.735
In	4.372
Sh	4.735

Appendix 10. Substitutions used in consonant production sorted based on FDHs' L1s